

Figure 1
SEQ ID No: 1

gtctataatggcagtcacacagggctctaaaactttgcagttttatcattaactcaaagtgaatgtatacatg
ccgctgactcaacattttgagagacaacaaatacaatgaatatcaagatacatatataatataatgtat
ctttttgagatggagtttctactgttgtgtccaggctggagtacaatagcacgatcttggctcactgcaacc
tctgcctcccaggttcaagcaattgtcctgcctcagcctcccaagtagctaggattacaggcatgtgccacc
acacctggctaattttgtattttttaagtagagatgggggtttcacctatgttggtcaggctgggtctcgaactc
ctgacctcaggtgatccacctgcctcagcatctcaaagtgtgggattacagggtgtgagccacccacccgg
ccatataatataatttttgagatggagtttactctgtcaccacaggctggagtgcaatggcttgatctcggc
tcattgcaacctctgcctcccagttcagatgattctcctgcctcagcctctcaagaagctgggattacagg
tgcagtcaccacatgcccactaatttttatattttcatcatgggggtttcacctatgttggccagggtgggtgtcg
aactcctgacctcaagtgtctgcctgccttcggcctcccaaagtgtgggattaccggcatgaaccaatac
gcttggcaataattttttaagaaaaaaaatttcaggttgcaacagcatccaaaaagtaaccaatgatttta
gggtgaaggggtgaagacaaatgtaaacctctttttttttttttgaaatggcgtcttgcctgtcgcctaggct
ggagtgcatgtgggtgcaattgcgactcactgcaacctccacctcctggactgaaacgattctcctgcctcagc
ctcccagtagctgggattataggctcgcgcgcgcacgccccgctaatttttgtgttttttagtagagacggg
gtttcacctatgttggccaggctgggtctcaactcttgacttcaagtgtccgctgccttggcctcccaaag
tgctgggattacaggcttgagccacccgggtgaaatgtaaattgttaaacctgtgttttgaaaatgcataag
tataggataagggagaattgactttctgaagaccagaacatttttagtcaatttcaaacacaatgtgagtcaa
ttgtataaaacaggttcttctcctgatgaggataagaatagtatccttgtcagatggaaatgccattcag
ctgtactttctagtgggttacgcccatagtagcactgttgatggaaccaggtatctgactttaggaaagatgt
tccccactggagctgacctcagaggagcctgaccaacttggggaaagttaaagatctcatcacgtggagaa
taggggaaggcaccaacacgtattgagtgtctactttgagcttaagggagaaggagaaaaggcagggaataa
acggaggatggaataagaataggtaattctccttaggtttaataataagtgcttgccataggaaggagccc
cagaacacagttatcaataatagagactcacacagagcattctacactagagctgctgtcctcttgaccaga
ataagggtaaggtgtgtgtgctgtccaggaaagttaggcagctaggaggtgatcagagcatacactactgcc
gccacaattctaagtgctctcccttaggggaatcctatttcttctcaggcacatttgtttattcattccatg
ttcactcttgttatttacttcttggcaggctttgtgttaagaattggggaaacaaggttgaataaaccaggt
ctgtaagaaaaggagctcacagctcggaggggcaaatgggcattgtgcctgcaagttggccactgagagc
ctaagaagtgaaagttatgaatccaggattactcagttatcaatgaagtgattaaacatcatccatacagacc
ttcagagctggagggaattttggatacctactcagcacatagttttcaaacagtgcttgtggaaccctagg
gcatttcttagggattgtgtgtgtgtgagagaggagattgaatcagaaggtgtctgggaccattctctactca
cacttcaagcagagcagctccacttctatctgtattattttttattttttatttttztattttttttga
gacggagtctgcctctgcgccaggctggagtgacgtggcacgatctcagctcactgcaacctctgcctcc
cgggttcaagagattctcctgcctcagcctcctgagtgtggaattataggcctatgccaccacgcccagc
taatttttgtatttttagtagagacgggggtttcacctatgttggtcaggctgggtctcgaactcctgacctgt
gatccaccggccttgcctcccaaagtgtgggattacaggcattagccaccatgccccgcctatctgtatt
atttattcattattgctatgtgaatgaacctgaagaatgcttactgttactgctaagtatthaaccacacc
catgcccattgcaggatgatagtgaatagtgggccaaaagatactataattagactcatgtaattaaggaatat
ttttgtcttgtacctattatgtgcctataaagactatgaaatctattttattcagtgtatttttggaaatacca
aataagcaaagatcctatgtgctaaagattctaatttgtgctaagattttccttcagatgtttggctttct
caaattccctgagggctagaactttgcctactcatttgtgtttcccaagtgtctaacgcagtgctgacac
atacaggatctccaaacgcttgcctgaatgtgtgaggaaggaattaaaataatgtaccgcccgggcaagtgcc
tcatgctataatcccagcactttgagagaccgaggtgggcagatcacttgaggtcaggagctcgagccag
cctggccaaacatgggtgaaaccccgctctctactaagaatacaaaaattagccgcaggtgtggcaggcgctg
taatcccagctactcggaaggctgaagtgggagaatcacttgaacctgggaagcagaggtcgtagtgcgcg
agatcagaccgctgagctccagcctgggtgacagagcaagactccatctcgaaaaataaaagtaaaataaaa
taatgtactaactggaccacagaacagattttccaattgattattgacaacaaaggaatctgaattatttaaat
aaggtgaataagtagcatattcatatatatatgtatatgtgtgtgtattttacattttttataaaagtgtaaaag
tatataacttttttttctttcttcaggtagaacctctcctagattgtcactgaataaacattagcactaa
ctatggcaatcaaatacatattgattgggtgtcagagaagaattgaacattcaactctgaagcagtggtatt
tcttccatctgcaaacactctgtccccatcchtcttcttgtgtatcctggaatccaagtcataaataatgat
aggatttctgtcaaaggtatttctcagaggagtgatatgtaactccctttcctctgatacactgactcact
aagcagctacccttgtgaattccaattagcaatacttcttgcctatgtctgggtccaactttcagacaaacct
gtgttcaggattcctatagccatttatagggtgtagacaggaagcattcagatatcccagaggtacctgata
accagctgatccatgactgtctgtcttgggcttggccagcttgaatatcttgacattgtgggttctczzccaga
gaaggtgccttttggatgtgagataaagacattatgactagatagtgcatgggtgaggggtgtttttctagt

[illegible]

gtcccagctactcgggaggctgaggcaggagaatggcatgaacccaggaggcgagcttgacagtgagctga
gatcacaccactcgcgtccagcctgggcccacagagcaagactccatctcaaaaaaaaaaaaaatcttttata
ctatatctttaatgtaccttttctatgttttagatacacaaataccactgtattacaattgcctactgtattc
agtacagtaacatgctatatgggttacgtagcctaggaacaataggctatatgttccctagggtatagggatgt
ggatccctataccatctagggtttgtgtaagtttattttatgatgtttgcatgacgacagagtcacctaagga
ctcatttcttagaatatattttagttagtaagcaatgcatgactctattgactcatgaattcttaccacagacc
tatggggcagtactattgttaccctcattttataaatgataaaactgagggtacagagacagtaaaacttg
accacggtcattcagctactcaagagtcagggtgggattttaaaaccagatcacatgggttcagagtgttca
cacttacctactatactgtctcaagagcaaggatgttttggttcacttgacaaatgaagatagggaacctctt
tcattataagcctatttttaggctaaaaatagaagggaaggacacagtgaaatccaggcctctggcatggct
cctcagccctttctgagctggcctgggacagccttccctactcactgatgccacttccctactgagcgacttt
cctgcctacctcactgatgccacttccctactgagcgactttccctgggctccagacccagtaagcgactttgc
ctgcaaccaccttattttgcctactccctgtgcttttatgcctttaccatctgccctggaaagctcttct
aacctttgaatgggtaaaggcataaatgtatgctggagaaatccctcagctcaggggccaggcacgctggctca
cgctgtaatcccagcactttgggaggccgaggtgggttagatcacctgaggtcgggagttggagaccagcct
gaccaacatggagaaacccctgtctactaaaaatacaaaattagctgggcgtgatggcacatgcctgtaa
tcccacagctactcgggaggctgaggcaggagaattgcttgaacctgggaggcgagggtgagtgagccaa
gattgcgctctgcactccagcctgggagacagagctagactctgcctcaaaaaaaaaaaaaaagaa
aaaaaaaaagaaaaaagaaatccctcagctgagttgtcaactcctctttgaaactttctcagaccttccaggc
tgagtcgctcgtcattttgtgcttccctcagttccctggcttctaccttcttcatagcttgtttcatgtaattgt
aattcttacttgcctttctccctcttctaaagctgagagctacttcaaagcatgggtaggacctagcacggtgt
atgggacatgggtggtaccccgtaaatgtttactgaaaaaaaaatgcctaaagcaattgttaacatcatcag
atagataattatgggcattcagagattctgtcttcaagcttatataaagaacttatttttggtctaatat
cctgataattttctcattactttcacttattgtggttggatcaattgttgacattttataaacatttca
ctatttgacaatgatgataactaaaaatcgaattaaagcaaccattctaaagatagtgatgataacatata
cgctggtaacatcttttattttcagccgtatcatggaatcctctgtttccattctgctaggtaggcaggtatg
caggtagaacttgtgagaggatattgatttttgtttccatcttagatatgacaggaacttggaaattttgaca
taaatgacgaacatccgggattcttaacaatctttaaaaatggaatgccttaaaagctgggcgcagtggt
cacgcttataatcccagcactttgggaggctgaggcaaatggatcacttgagttcaggagttcaagaccagc
ctggccaacatggtgaaaccccatctctactacaaatacaaatattagccgggcgtagtgccagggcgctgt
aatcccagctacttgggagcctgaggcaggggaattgcttgaaccaggaggccttggagattgagtgagc
tgagactgcgccattgcactccagcctgggcaacaagagtgaactccatctccggaaaaaaaaaaaaaa
aaaaggaatgcctttgggaataatattttataatattatgtataacatataagacaaaccattagtttgtctt
atattttactaaatataaatatttagtaaatataaatattttactaaatataaaaaactcttagattttactaaag
agttacaactaattggcctggcgtgggtggctcacacctgtaatcccagcactttaggaggcagaggtgggccc
gatcacgaggtcaggagatcaagaccatcctggctaacacggtgaaactctgtctctactaaaaaaaaaaaa
atacaaaaaattagccgggcgtgggtggcagggccctgtagtccagctactcaggaggctgaggcaggagaa
tggcgtgaactaagcagaggcttccctaaaagtgtcttcaggataaaggcagaggaagaggctccatgact
gggattgggtgtgaggagagccagagaagcaagctacagaaaaagagaaaaaattaatatgcaagagagtaaa
caacacgaaggaaagaacccagtggtggaacactacagctgagaaagggtgtctgtaaggatgttctacaaa
gcaaatgcttggatataattcattgcagcaggagatggtaagcctcatgataaagaaggagaaaaaatcaa
gtcaagggtctgagggtactgaccaggtatacttgactatgccagcaactgtttagggggagatttgagct
acacttgtagcaaaaggcaaaatctgtaattagttgtaactcttttttttttttgagatgggtgtctcgctctgt
ccccagggtggagtgagtggtgtgatcttggctcactgcaagctccgctcctgggttcaagtgtattctc
cagcctcagcctcccaagtaattgggactacaggcatgcaccaccatgccagctaattttttgtacttttat
tacagaccatgttttgccatgttcaccaagctgggtctcaagctcctgacctcaagtgtatcgctccgctcgg
cczccccaaagcgtgagattataggcctgaaccaccgcgctggcctaaagagatctaattcttagcaaaag
tttcaccagggtgtctctcctcaccocccaccccatccttcccacaaagaattagaacaatgtccctactacc
cctgctgtatctctgactttttactttaaatctcagcagaatattttactaaatgttttgatgtgggtatat
aaaatcatccctgctgacaaggaaacactttttgaaaaagttttcattatcaaacagtaagtacagctgac
tgccgtgacctttaaccattttctgagttccctcattggacttgggtggaggggactgggtaccaataaag
tcaaatgcttaataatatttgcaagtgttgaagaaatttgaagttgaatattttctatcatcttgaaatgga
gaaagaatctgtaaacagcaaaagccagacgcctaaaggaaaagatttacagattaaaataagattgcaatc
tggtaaaaaattttgcaacacatgtaacagtcagaaaagttgaaacacttgggttaacawgagcttttawcag
ataaataaggaaagaataaacatttggattttaaacaccgataaacatgaaaagatgttttaactctttatttt
atthaatccatattatttttcagtttaatacaagaaaataataatcaaaacaataatacatttttatatatat
atatatatatatatatatatatatatatatatagtaggcataaggttaagactgataagactgttgaaaagg

gtagtaaaaaactaggcttactctacaccaaataatatatatctattaggatggctaaagtaaaaaaactgaaaaatat
caagtgtctaaaaaggatattggagcaatttgaaccctcagacatcgctgtrtztgagaaaaacaaaatggtagcgc
caccctggagaacagtttagctgtttcttgtaaagttaaacatgcgcttaccatatgactcagcaatctcac
tcttgggtatattatgctaggaaaaaggaaaaatttatacttgcacacaaaaaacttgtaagtgaatctttatag
cagctctattcataactgccaaaaactgagagaaaaatgtcctttaatgtgtgaatggataaaccaactgtgc
aacatccatgtaatgaaatactacttagcaataataataataattataaaaaacccagaaccattgatgcat
gcaacaaatatggataaatctcaaaagcattatgctgagttaaagaagtgcgtctgaaggatttcatactct
aggattccatttatataacattattgaaacgacaaaattatggggacagagaatagatcagcgggttgccagg
ggtttaggtgtgtggagagggtgtggctataaagaacatgcaaggggaattttttggggagatgaaatggatc
tgtatcctaattatgggtcatggtaacacaaaatctatacatgtgttttagattcatagaactgtataccaaaa
aaaaaaagtcatttttactctcttaaaatgaaaaaagaaaaagcctgggcattctaacaccttgtttgtgag
agtacacattgataccaagttttatgggtgggcaattttgtctataaatacggaaagtttgtctgttctattat
tcagcaatcccagttttgcaaaaactgtctaaagaattcttggggcgccgcagcgggtgttcacgcctgtaa
tctcagcactttggggaggccgaggtggacgatccctgaggtcaggagtttgagaccagcctggccaacatt
gggcaaccctgtctctactgaaaaatacaaaaatttagccgggcatgggtggcgcatgcctgtagtccgagctac
tcggggaggctgaggcaggagaatcacttgaacctggggaggcagaggttgtagtgaactgagatcgtgccacc
gcactccagcctgggcaacagagtgagaatccgtctcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacaaaaa
aaaaaactttgtgtacgtgtgcaaagagaatacaaatgatcatggctgcattttttaaatgactataaaaa
agaggtacaaccagccaggtaaagtgggtgtgcacctgtagtccagctactcgggagggtgaggtgagagga
acacttgagtccaggagtttcaggccagcctgggcaacatagtgcagccctgtcccaaaaaacaaaaa
caccaaatgtctatctgttaggaatttgttttcaagttgtgatacataggtacagtgaatattatacattc
atttaaaatgatgataaaatctgtattttgtttacatgaaaaactgtccactataatgtagtgaataatag
attacaacaatatatatggaataaacttgttttagaacaatttctagaagaaggtagaatggacagaatta
tctctgggaagtgggtttataatgattctcattttcttctttgtatctttttcatagtctttctacttttgt
tatgtctggacatttgattatgagcatgtattactgatctattttaaaaaattgattttaatttttacaaaa
ctcatgtaaaagcttgaaggttcgcatttttagaccatgttaaaatttctctggatcaaaacagacttattcaa
atatcttgtaccctgtctccaaaaattgcctgccccaaaaatacaactacaaaagagagcatttagctgcata
tttggactgtctgagatcaacaattattttaccatggcttaaatttttaccctcagatgtgtgtgggttaca
aacactctccacatttttagaggcattgtcttttgatatttttaaatgtaaattcagctgtgcggtggctaa
cgctataatcccagcactttgggacgctgaggaaaggatcacttgaggtcaggagtttgagaccagtttag
ctaactgggtgaaaccccgctctcttttaaaactacaaaaatttaaccgggcatgggtggcaggcacctgtaac
ccagctactcaggaggctgaggcaggagaatcacttgaacttgggagacagaggttgcagtgcagccgagatc
atgccactgccctccagcctggccacagagcgacactccatctcaaaaaaaaaaaaaaaaaaaaaaaaaaag
gccaggcgagtggtcacgcctgtaatcccagcactttkggagcgccaaggtggssggatcaactgaggtt
gggagttcacgaccagcctgaccaacatgcagaaacccygtctctactaaaaatacaaaattagccgggtgt
ggtggtacatgcctgtaatcccagctactcgggaggctgaggcaggagaattgcttgaacccaggaggtgaa
ggttgtgttgagctgagatcccgccattgcactccagcctgggcaacgagcaaaaactctgtctcaaaaaaac
cgaaaaaattcccccaaaaaaccccccaaaaaacagcaacaacaaaaaatacaataatgtaccttgttttagc
ataaagcataattatatgcatatgggtgattgggaggatgaaatggaaaggttattttactgacttcagaa
attatgtcctgatagatttgattgggtgatttaaatataacttcttgtcaagcatctgttttagaatcaaat
actatgactctgcagtttcttgaactctcatagtatcacatctctgtttgcctttgcattgggttttaagaaa
atgaggaggtgtgaacacttcaacttcgtttcatgtattacatttttgaatgacacactggctatttctcta
gaaataaaggtgacaaatatttcacagaaacataagggtgctattatctcattcaatcttaggtcactcaaa
actctttctctccacacattgaagattcatttgggaatgcttttgtcttatttgtcacccccagtgagggt
gtggtaagtgtttttcattttgtcttctttgttttatctacagggttccattcaataaaacaaagggacttggg
tcaaaactcaggctcttatgggtttggatgtaatctttgggtctcatttttagttaccaacagagaggtgtgc
ttctgacctctttgactcttccctgtgtaatttactatgcctttgatacttgtgaagggtgagattttcgag
gagtactgttgtttttgttagaggttgtaatgtctttcttctgctttgtgattcaagttgtgttcagttacaa
tcataagcatgtgcctaaaaaaatcagatgcaaaactagcaaaagtagaaactcaggggtgacagctcttaaga
aaagatgcaattcttggggctgggtgcggtggcttatgcctataatcccagcatgttgggaggccaaaggtgg
gcagatcgcccagggtcaggagttcgagaccagcctggccaatgtggtgaaacccctgtctctactaaaaata
caaaaatttagctgggcgtgctgggtgggtgcctgtaatcccagctactcaggaggtgaggcaggagaattgc
tggaaactgggaggtggaggttgcagtgcagcctagattgcgccattgcactccagcctgggcaacaagagcg
aaactctgtctcaaaaaaaaaaaaaaaaaaaaaaagatgcaactcttattactgacacagaataaaaaatttag
ttacatagatttgtaaaaggactatcagctaggtttagccttaccagatttaggtaaatcattctctgtcta
cactcatattctcagccacttccctcatcacattttcagggtgcagtatataatagcgtcaactcgtgtaat
tccccctactccccatgaacttctagqccaaqqqqccacacqqqqttqqqqcatataqtataaaaggagtaagq

[illegible]

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

taatagaaaacagtatctggaacattaaacgtaggatttttttttttctaatgacttactctcttattaat
atgtcagagaaaagaatagctcctggctaagaaatacaacagtcctcatcccagaaatcacagccaggaat
atggattcttaagtgtcaaaaaagtagctctgaataaggaaagaaacgcagatgcactacttctattataata
agcgcttgatttcttttaatccccctgagtcgaattattttcttgccaaatttaagggtactgactgcttctg
tgaatctattgttacacttgataatggatctgagttgggggtaataatttgcctatcaatttggatacttaaa
aatctctctctctttcttcatataacctctatctcacaactttccatttaatgaggggaagtgaatttctttt
ttctgccccctctttcctactgcttctagaataaaaagcataaacaggacgacaggagtggagatgagaggaagc
atttccaagcaatgggaaagtatgatgagagtcgtgagttggtagaatgggggtgagtaaggggtgggggtgt
ggatggatggaagaggatggaggaggaagcaggtcatatgatcaggcttcaaaggcctccaatatcttgtt
ttcagaggcctttgtaggctcttcttacagaactttgaggggtatttttcttcttgtagcaaaggagaaccatag
taaattttgagggatgggctgggcgcggtggctcacgcctgtaatcccagcactgtggcaggcagaggcagg
tggatcatttgagggtcaggagttcgagaccagcctggccaaaatgggtgaaaccccgctctctactgaaaaaac
aaaaagttagttggatgtgggtgcatgtgcctataactccagctactcgggcagctgaggcagcagacttgc
tggaaacccgggaggcagatgttgcaatgagttgagattgtgccactgcactccagctcgggagacagagtga
gactccatctcaaaaaaaaagaaatttttttttttggagggagaaaatatatgattagattgtttttgtgtt
ggtttgcctgtttgttttttttcccccaagcaaaaaatcactttactgcaatagggaagacaaataggaagggg
aagaaactagagacaggaacagcagtttggaggttctgcaatacagaagccagatgtttggcttggactta
gacactgggaatgaaaaataaatgatgaattaaaaaaataaaatatttgggaggtatacttgacctgacctt
gggtgctttttcaaatagagaagagaaagatagagacagatgaaaaagttaagcaaggatgactatgatttccc
atagtgaatgtctggatgatgatcattaaatgaaaatttttaaaaaggcagatggtagaaggagatttga
aggaaagacaagaaatttgtttgttttggatttacttgtagaatgacctgcaaaattttgtcctaataaatg
ttaacaagtggctttccatacaaaaaaccaaaccacaaacccctgatgtaaacaaaaataaaattctgag
gcccccttcaccatctgaatgaacttctcctctgcaagggcactcttaaaatttaacatgaaagactgggt
tcaggtcatgacgggaagtgggggtcggaaggcctcattatgcctctctggcattaacatcaacacagacc
ttaagtctgttaagaagcatttacaatctattctctctgaagcctgctacctgaaggcttctctgcacact
gagaactttggctctccacaatcctttatcttaagccagacatttcttctatttgatcccagggtctttagat
aaactcaaccaattgtcaaccagaaaaattttaaatctatctataacctagaagccccacttcaagttgcc
ctgcctttttgaactgaaccaatgtatttcttaacttatttgattgaagtctcatatctcctaaaaacca
agctgcacccccaccaccttgggtgcatgttcttaggatctcctgagggctgtctcctgagggccaagatca
ctcatatttggctcaccataaatctctaaatattttacagagttttactcttttcatcgacactgatttata
ttggattttcaaatgggtgtaatatctcatgccttggccgatttcaagctagctggtagagatatgcaaacaaa
caaac
tgatatttctaccaggcagcataggggcgcaggtgggtgagcctggcaaggagacaagagctggaggtacag
ggcttaaatgaagccaagggcgtgaataatattgatgaggagagcaagtagaaatgaaaagagaaaaccagt
attaggagatctataaagaaacaatagtcaaaaaagagtggttaagaggcagaagaatcagaagagaacc
agattcatgaaataaaagagaggaatcagtttcaggaaggaaagtgtggtcagtagtgtcaaataccasaga
ctgaatagytcaaggattttaaagaggtcaacagcwcaccaaattaagcgggttttctgagcttaagatatt
tttctttatttttctatgttataaaatattttataaagaaaaagaaaaaagagatcacatttgccccccaa
ccgctccctggcctctctctctcttaaggaagtcgttaataaaaggggaggggtagagatgaaatctttccagagt
gtgaaaatttgactgcagtgaggtgaggtgttaataaaaggggaggggttagagatgaaatctttccagagt
tggtcttaaggggatagcttgaaggaaagaggggttgggaaggggtgggtttttaaagaattatcactcacta
atcaactagaaatccagtggaatatgcagtagctgtctgtaatccagcaggctaaactttttttttttttt
ttttgagacggagtctcgctctgtcaccacaggtggagtgcagtggtatgatcttggcttaccgcaacctct
gcctcccagttcaagcaattctcttgcctcagcctcccaagtagctggaactacaggcatgcaccaccatgc
atggctaaatttttttgaatttttagtagagatgagtttcccatgttgggtcagggtggtctcaaatcct
gacctcaagtgatccccgcctcaacctcccaagtgctgggattacaggccagagccacagtgcccagcca
aggctaacctcttgatcccaatgacaaacagaacaaacatcttactcaagtcagaagcaataataatttg
aatcttgcttgcatgtcaacaggagccacattaatacagaagaggatcacattgggtccaattaaattgaatt
gattgagagcctctgcaatacacgggtctactgcacaaataatgatggttcctgggtacattttttatttgacca
ttgattgctcgatttgtttctgtgtctaaactgtgtattggaattaagctgactcaatttgaactgcaggtcc
ttttatccctcttattttttattttattttatttttttaaggcatttctcactctccagaaaatctct
aagatttcagctattaggcatttgtcttt
ttgcccaggctggaatgcagtggtcacaactctgggtcatcacaacctccacctccctgattcaagtgattct
cctgcctcagactcccgagtagctgtgactataggtgcacaaaccacagccagctaaattttgtattttt
agttagagacaggttttctactgttggccaagctgggtctcaaaactcctaactcaagtgatctgcccgcctt
ggcctcccaagtgctgggattataggcatgagccacgggtgccccaccagggtctcttaatacaacaaatt
catcttaaaacatcattttaaataatatttttttttattcttcattgacttctgcattactctattttttt

agagtttctagcaaccagtatcatttggtatttttaacaatgtgtacatgtacatttatgcagatgagttaac
atatatcaaagcaacctccaaacaatgccatttaggtaatctccaatttaaagcctcaatagaatgataag
attgagcttttctgtagttccatgacctccagcagagtctgcaaggccacagctgcctgaagggttgattctg
taattagaagatgccaggggtcatctcagaatagaacctcaagccacccaggctacatttacagaatcagcct
ctccagaaaaacagcaacaaaggagggccttccatgtatttgggaaggagtcacctagaggagggacttggg
gttttggtgttggtgtggggggcagggatgggatggggagggggaagcttattgaaatatactaaaagacaa
accaacctaaagggtggaggggaagaaaattcacacttgaagcttctttttaaggggcatctcttagggtc
tagcttttgagattcagtatatatatttttgagtcttgctctgctggagtgcagtgggtgtgatctgggct
cactgcaagttctgcctctcaggttcacaccattcttgctcagcctcccagcagctgggactacagggc
gcctgccaccatgccagctaatttttttgatttttagtagagacggggttccacctgttagccaggatgg
tctcgatctcctgatctcatgatccgcgcctcgccctcccaaatgctgggattacaggcgtgagccacc
gcgcccggccagttatttttggtttatgaagatattacatttgtaagtatgagcttggtgtcagcaaaact
atatccctgtgtacaaaactggccgaatcacttagccactttgggccaatcacttagctcttctaacagtaa
gaaatcaacaagaaaaataaacatttcaaacattacaatgtgttcattgtattcactgtgggggatgaccaga
ttctcgaaccacaggttgttcttagtgaaacaagttgggttggggccatagacttgtgtatttagaatca
atggctgttctctctctgggactttgattttttcttgggctcatcccttttttgtagtatcttattcttgt
cttatttgtataggacttaactgttccattcccttatttagagcaatctaagtgattaccttcataaccttt
ggaattatatgcttcaaaattccaaaaagaatgattttgggctgggcacagtggtcacacctataattcca
gcactttgagaggctgaggtggatgcctgaggtcaggagtttgasrccawcctggccaacgatagtgaac
ccgctctctactaaaaaatataaaaaattagccgggcatggtggcaggtgcctgtaatctcagctactcggg
aggtggaggttgcaatgagccagatcgaccattgcactccagcctgggcaacaaaagggtgaaactccatc
tcaattaaaaaaaaataatgattttgggtgctgacttcaaataggtaggagaagaaggagagaggagatgg
agggctcasggagatctaattactctctaaaatcatgctaggaaagataaacaccttttaataacactctctgc
ttttataacatcattctgccaaggagctcaaaggtttcaacamagttcactttcagaaaaacccctttgagga
agacagaatatacatcttctctccmttttaagatgaagaaacaggccgggcacaaatggctaagtgcctgtaa
tcccagcactttkggaggctgaggccasargatcgttgagctccaragtttgagaccagcctggataaacat
ggcaaaacccctgtctctacaaaaaaaatacgaataatagatgggtgtgggtggcatgcacctgtggtcccagc
tacttgggaggctaaagggtgggaggatcgcttgagccaggagtcaggtctacactgagccatgattggatc
actgcactccagcctgggtagacagagcaagacctgtctcaacaaaaatgaatgaaagagaaagaaagaaa
gagttagaggagaggagatgaggggaggggagggtagcagggagggggggaggaaggaaggaaggaaggaag
gaaaaaagatgaaaaaagaaatacgcaacatgaaacagaggcagaaagactttacgtaaatgctcatcat
gtggttgtaagtttgacccccaaaacccaatttattgaccaaggttattctttgactgaggcaagggggtcc
gctctcctgggcttgggctttagaaagctcatctctggcctttctgagatccatccctttctttttatttt
tcttgacacggagtcttgctctgtcactcaggtggagtgcagtggcatgatctcgactcactgtaacctct
gcctcccgggttcaagcgattctcctgcctcagcctcctgagataacaggcgctcgccaccacatctggcta
atttttgtatttttagtaaaagactgggtttcatcatgttggccaggttgggttcgaactcctgacctgaggt
gagctgcccaccttggcctccaaagtgcctgggattacaggcatgagccactgcgcccagctcagatccatc
cctttctaaggggcaaacagtcctggtgcaaaggggcatgccaccagagttatgagtacctgggactcca
gaattccttgccctgggtggcctccacatgcacttccagggcctgcttgggctcttctatgggtctgtcctga
gtgttgatagaaccactgatgtgagtacctgggcttgagccgtggcctggagatcctgttgactgtagcatg
gagggggcttgtgcagctgaatgtctgyatgcaggtgggtgggagttctggaatatgatggagctggaggtgg
gaagagaagtaggcttggggcagctctctcatgccacctcattctggccaaaactcaggtcaaactgtgaag
agtctaaatgtgaatctgccttcaaggtggctacaaaggtatctttgtcaaggtaggagaccttgtggcct
ccagctgcacttccagggcctgcttgggctcttctacgggtctgtcctgagtcttctatgzaatctgtcct
tcagggcagattcataatttagactcttcacagtttgacctgagttttggccagaataaggtgacatttagtt
tgttggcttgatggatgacttaaatatttagacatatggtgtgtaggcctgcattcctactcttgccttttt
ttttgccccctcagtggttttgggtagttttgtctccctacagccaaaggcaaacagakaagttggaggtctg
gagtggtacataattttacacgactgcaattctctggctgcacttcacaaatgtatacaaaactaaatacaa
gtcctgtgtttttatcacaggagggtgatcaatataatgaaattaaaagggggctgggtccatattgttctg
tgtttttgtttgtttgtttctttzzztzzztzzztgtttttgtggcctccttccctcctcaatttatgaagag
aagcagtaagatgttctctcgggtcctctgagggacctggggagctcaggctgggaatctccaaggcagta
ggtcgccatcaaaaaatcaaagtcagggttgtgggggggaaaaacaaaagcagccattaccagaggactgt
ccgcttccccctcaccacagcctaggcctttgaaaggaaacaaaagacaaagacaaaatgattggcgtcctga
gggagattcagcctagagctctctctcccccaatccctccctcggctgaggaaactaacaagggaaaaaaa
aattgcggaaagcaggatttagaggaagcaaatccactgggtgccttgggtgcccgggaacgtggactagag
agtctgcggcgagccccgagccagccttcccgcgcgtcttaggcggcgggcccgggcgggggaagggg
acgcagaccgaggaccttaagacacctgctgtacctccazzzczzccccacccacccacctcccccaac

tccctagatgtgtcgtgggcggtgaacgtgcgccgtttaagggcgggcccggtccacgtgctttctgc
 tgagtgaactacataaacagaggccgggaacggggcgaggaggagAGAGCACAGGCTTTGACCGAT
 AGTAACCTCTGCGCTCGGTGCAGCCGAATCTATAAAAGGAACTAGTCCCGGCAAAACCCCGTAATTGCGAG
 CGAGAGTGAGTGGGGCCGGGACCCGAGAGCCGAGCCGACCTTCTCTCCCGGGCTGCGGCAGGGCAGGGCG
 GGGAGCTCCGCGCACCAACAGAGCCGGTTCTCAGGGCGCTTTGCTCCTTGTTTTTCCCGGTTCTGTTTTT
 TCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGAGAAAGAGACGCAMACACAAAAGTGAAAAACAGgt aaga
 ggctctccagtgaacttacttgggcttattgttttgtttcgaggccaaggaggcttcgggaagtgtcggtt
 tcggggactttgatccggagccccacatccccaccacttgcaactcagatgggacccggaggcggtgttaaat
 ggggagacgatgtcctagtacgagctctgggtgacccaggactctgcgctgctgcgcttgzgggcttgccccg
 acggtggagaccgggagcatctctgggctggagacccggcgagtagcccggttcagaggggtcgggg
 gttcccggtgctgagggcgctgctgcgggtgggagagctgcaggtccggcacccagzcgctgcttt
 gttcggagggcctgagctggcztagzaaaccttctggttgaggctcgccagtagctacacggagacaaatg
 ccagcactgagttctactcggttcttaagaagctggtctgttctgacctgggaattggctatatgtcccc
 gggactggagcggcacagttccggactgtgaatccgggaactcgagttggagggtgtcccaaaggctcggtgg
 tgctattgtcactagaggccttgggtctttgtzttgacctgaggggtaggaggtctgctacagttctcc
 gtgcgctcagctgagctggtgtccctggcgagagcgaggacaggttttgtttctttttctttttcttttt
 ttctttcttttaagtctcggtctgtgcgccaggctggagtgcaatggaacgzatctccgctcactgcaacct
 ccgctcccggttcaagcgattctcctgcctcagcctcctgagtagctgggattacaggcgctcaccaca
 tccagctaattttttagttagagacggggtttcaacttgttggccaggctggtctcgaacctcga
 cctcaggtgatccacgzgctcggtcctcccaagtgtgggttataggcgtgagccaccgcgcgggcg
 agttttgtttcttttaaaacaagacttaggagagcctgaggagaccggagggtgggtgccaatcctccc
 tctcccacgttccctgcagccccatctccagaccgttgcgtgctggtctctcgggcagcttctgcctgggc
 gcagatggggaagctgggcccagggtggtgctggaatgaccgggagtaacccggcgggcggcagaaact
 cggagctccgcccgggggtgggtggtgctgctgaggggtgggggtgctgggcccggggctgaggtgg
 zccccggagactggccgcagzcgctcctggccggaggacactaggaatcgccggctctactaggtgtctt
 tgctcgcggttcgactgtgaatccgggtgaagaccgggtggtgagacggggaggaaactatgaggtgaggc
 gaaagccggttttgtttttttttttttgttttttttgggttttttttttgttagcgtgtttgccaaactccag
 gccattggttaaagcaggaaggttcttggggcgggcggacgggtgcccagggttatgtgtagggtgctcttaggt
 atatcttttatcaaaaagaagcaaaagaaataagattaaaaataaacaagaaaaaagtgtctggcactggc
 agtaattggcctgcttttgcagcactgataccattagcttttaaaatccgacttttcattgacacttcaaga
 agagaatgggttagtatatacacattcatctcatagtggacaaatttcatatttaaaaaaaccttctgggtac
 tgaaatcagcaagtcacttgccctccatggccgaatccctgcttcccacgaagagaacctcacaaaaatttc
 cccaagttaaagagtggaaattttcttgatttttttztcttttttttttaacggcgtagtttagaazccc
 agacttaattatgatcttcttttcaacaaaaacttaagtccttaagttttcatctccccttttatttcaa
 cctattcttctcatacctaccacaaaaataatggaggctttctgttgagaaactttccgtttctgttgagag
 tatcattctcttgagaaactttctcctaaatcagagaaagtatggaagcatggaaggtattcctgagtagaa
 cctctacagatattacaatatttttcaaatacaaagtttccattgtcagcctgtttcccaagtgttccaca
 aaccattaaataattccacaaaccattaaaataattaatgctagggaaatttttaggaaaacattgggtttacaa
 tcagaaggaccggggaagtgggtcttcagccttcacgatgactacaagccatttaagggacatagaattgtcta
 ctgttgtcagagcaatttaggagttgtatttgagaccgcagatagtttccagaatgacatatctgactgt
 aacctggacagtgatgtgtgtctccctgcagatgagcatttgaaatctcaacctcgtattttctacg
 agtgcaggcctataatggaccctgggcacatttttttttttttgagatgcagttctgcgtctgtttcccagg
 ctggagtgcagtggcagatattggctcactgcaacctccacctcctgggttcaagagattctcctgcctcag
 cctcctgagtagctgggactataggcgccagccaccatgctggctagtttttgcattttcagtagagacag
 agtttccaccatgttggccaggatggtctcgatttctgacctcttgatccaccgcctcgccctcccaaagt
 gctgggattacaggcgtgaggcacccgcgccgacccctggacacattttgacttagaacatattttcggttt
 gtgtgagacagtgcattagtgaggattggaaaagagtgatcaggaattgattgtttcaaggattgggtcc
 ttctgctcaaggaagtccattgtaaacataaaaaaatgaatgaaactgaagaagttcagtgacttagcttt
 ttattatctctgtagtacttaccttttggagaggagttgggtgggatatttttccatttaaatttttttt
 ttaaagggtatctctctcccgtaagccgggatacttaagctatatatgtagtggctacaaattaagggtcttc
 actgttttcattttttagctgctagaataagtgaacattaccttagatagactcttctaattatgaagatctc
 tagatgtctagaaaatatcaaaatgcatgtggtttttgcattttctaaaatacttttaaaacaaatactttt
 tctttttttttttttctgagatggagtcttgccttttgcctaggctggagcgcagtagaatgatcttggc
 tcaactgcaactccgcctctcaggttcaggtgatttctcctgcctcaacctcctgagtagctgggattacagg
 tgctgtccaccacacccggctagtttttgggttttagtagggacagggtttcaccatgttggccaggctgg
 tctcaagctcctgacctcaagtgatctgcagcctcagcctcccaaagtgctgggattacaggcatgagcca
 ccacacctggcctcaaatacatttttttaagtatccagatatataaataaataataaccattatagtagttgtt

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted March 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

atgggtcatttactctagcatcaaagtgttaaaagatcattctgaacacttggtttgtttatgctgagagaagg
cctactccaaaaaatgcaaccatcttcgtatctgcatgtggatacaaccatgaatggccaaagtatttgcag
tggtgaatagacacttatatagcactgtgtggcaagtactggttgaaatgttttccagtggtttattcattg
tattttattttgagataggggtcttgctctgttgacaggggtggagtgacagctgcacagacaagggtcactgca
acctcagcctcctgcgctcacgtgatccctatacatcagcctcttaaggagctggggccacaggcacgcacc
actgctcctgggtcaaattttttacaattttttgttaaagacaagggtctcactatcttgctcagactgggtcttga
actcctggggtcaagtgatccctccacatcagcctcccaaagtgtgggattacaggcataagccactgtgc
agggtcatcaaagtctaattgaattttcacacaacaaaccattttattgtccctagttttacaagattaaagtaaat
gagaagctaatttttctctggctatataccttgcaaggagcagagctaagacttgaaccagccagaggttct
ttaactccagcactaacatttcagctgctgcaaccagggagcttttcaaggatgatcaccacattctctaca
ttcatctgctataatccttatcagaatctacagcctgtatcatattttccttggtgctgtgagtggggtcagc
caaattctcttttaacttgaaaccttggttgctgtagggattgcaacatcctggaaagaatagaataaaattta
ctcaactcaatttttttacttggttcataatgaaaactatactattgcttcagtcagatgtttgcgaatagct
gtgtgatctcaaaatgttttctatgtgatctatagzaaaatggaatgatagagtattagggtgtaagggcc
taagazaacaaaggaaaaagagaagtgaactgttagtttagttgtaaaaccttaactttgggtgaattgtaaaa
atttgtttataatacaatatgattcttgcttgctcctgtccttgatgaagttgtggaccttttgaaataagcta
tttctctgttactgctgttactgttttagaaatcaaatttagtttttcttaagatatacgtatttttggaag
ataaacacagtttcaaagtctgccttggttggtggtgactgggtcattgttgattcccagcactttggg
aggccaaagcaggaggatcacttgagggtcaggzagttzcaaagaczagzcctggcaaatatgggtgaaaccc
cgtctctactaacaatacaaaaattagctgggctggtggtggtggtgcttgtaatcccagctactgggattgg
gaggtgagtagaagaattgcttgaacctgggaggcggaggttgactgagtcgagatcgtgccactttac
tccaacctgggcgacagagtgcagactccgtcttgaaaaaaaatgtctgccttgtaaaagtgaataggatga
gaaagtgcctttcttatttaattgggtgtaatgaattagaaataaactctttgaagacacctcttggtaaaaat
agttacatttactgttgatttatgggtatgttggtatgtttttagttttccgtgtaataactcagttcatt
ctcatgagtgaataaggtgcttttattgtctttatagatgggaaactgaggtataggcaggttaaggtacatt
attatggagttcgttaagttagtggaagtcgaagtcgcatccagacagtttggtctccgtgagtttaccatct
catggtaagactttgtcagactatcaaagttttgacaaatgaaatattagcaaaaggccaaaagggttct
ctattttcatttgagtatcttcacctgaaaatagttgctgaataagtagcctgcatagaagggtacatttt
agaaatacttgaggccagagaatgaaaagcttacataaaaattgattttccggtggggccttcagttactctcc
attctacgaagaccacaaatagcattcaggcaaaagagcattttatccaacaatggaggagcactggatttgg
ttcctaaaaacaaaataaagtttgaaatcctgtctttcccatgttgaaaacaaagtgggtacaaaaccttta
gcttttgcaaacctcctttaagacccgatttaaatgcytccctcctcatgaagctcttctggatccactcyt
tcccatcactaagttgaaagtaagatccccctctctttacttccattagacttggattacagcactctttgt
atcatgtattttaattctgttttttaattacagtttaacattttatttgtcttctcttgagtgtatgcttctct
agaggaagggtctttgattcattctccccctggccttaattcatcccacttaatatggaaaaaatttaataaat
gctgacttgaataagttccaacaaggagaatgggaagctcatgtttgcttctctctctctctctctctctct
agataacagggttaatcacagaaaagcattagaaatagagttatatgagaaacaactgtagtttaaggctaggt
ttatgttagactgagaaattttagtgcatacttaagttatttagggccaggttactttttagtagaacaacat
ttcagtttgcgtcagtttcatctccgtttctgzaggcagctgtgatttaagaaaatgctctagctctgtggca
ttccatattcaagtaactttgagttgtatattatattttttagttaataagagtgcactgactcactaagta
atttagagatttaaacacttttttaaaaaacagtaacttcatatgcattggatctattctctataaagctct
ttcttgggggggtgtttgttttaaaattccccgggtgttttctctgccaatccaacttccaagaagcatttgg
aagtcaaaaacattttatctggttagtcttaaagttccagatattttgtgatagctgggtatttagtttatgata
tttcccaggaagaacttttttagtagttgaacctttatgaaagacttccctgaaagctaccttagagagttga
tttagttcttctctaaataagtaaatagaatattagttataggacatcttgaggtatagatgcaaatattgggt
gaaaaagaacatggatatacagagtcaaattaatgtagattggaattctgggttattactaatggatatctgac
attaggcaagttgctgatcactctttgcctcagtttcatcatctgtaaaatagggtatttgtgtttgtgtata
atgtgaaccgtataataatgcttggcctatagtgaaatttattcatacagagtgttttcagtgattttaaa
agcttgcttttagggcggggcgagtggtctctgcctataattccagcactttgggaggccaaagggtgggaggat
catgaggtcaggagttcgagaccagcctgaccagcatggtgaaaccccgctctctactaaaaatacaaaaatt
agctgggctggtggtgcaagcctgtaatcccagctactcaggagggttaggcagaagaatcgcttgaaccc
aggaggcagagattgcagtgagccgagatggtgccactgcacagagcgagactccatctcaaaaaacaaaa
caaaacaaaaacaaaaatcttgctttatagtttacttccacatcaaattgtctttatcccatggttacttgc
tgatatcccagacatgaaaagaaaaaagatgataacaatgacagttattaaattagggtccactcttattc
tagatccaattcatattactattcagacttggaaacattaaatttttagttaaactttttcaaatatgca
tataattgtcagtggttactataattttggggaagagattgttgacttctttgaagaaagatacggattttct
cttcagaaraaatacacatggggtcatataatccaaattttatgtgtaattacagggtgttcatgaatgccc

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

acaaatccattaagccatgtgaccttggacaagtcattttacttttctgttttttaggttgttggctctgttaa
aatgatactacttgacttttaagagcccttcaagctcttatgtcctctaaccacaggtctgtattcagaag
aaggggtggctctttaattagagccatctagagatctgaggaacatgctgggcatttagtgtaacataccatg
tggattttgagaggttaaagaaaaataaccaggggaatgcctcagagcattcctgatcagatcgatgacagaa
gaaaggaatgagaggggagagaggaagctgttgaaatttccatattacctgctttgagtgaatgaagatttg
aatcatagaaccagaaggggttctcatctgaaatgcaaaggaaggaggaggttgggttaattcaataagtttc
agttgagtaaacatgatttagtgagatactgttcttgcctctgactcaccatttggaaaatctctctaaaat
aaaattggactctccatctcggacatcattttgggtgtaggttttgcttttttttttgagatggagctcgt
tatgttggccaggctggagtgagtgaggcgcaactcggctcactgcaactccgcctccagattcaagcag
ttctcctgcctcaacctcctgagtagctgggactacaggcgtagtgcacacatgtccggctaaattttgtat
ttttttaatagagacgggttttactatgggtggctaggctgggtctttaactcctgaccttgtgatctgccc
ccttggcttcccagagtgctgggattacagatgtgagccacagtgcccgccctaagttttacttcttataat
ggactcctgttaagccaataggtgatgaaaggaaaccataaccaactcttcaggctcattcatccttcaaga
atagcatgctagtaccatcctaggaggagaattggactatacctcatgaggatagtttgaagtatctcaga
agaccctcactgggggtaggctcggttaagacacaaagctttctaaagcactgtaccaaatttgttgtttgaga
gatcataacaaattagaagtggaaagaagaaggagtaaaaggaagaagaggtttctggccaggagcagggag
ggggaaggagctgctaggaagatgtttgggtgtcatatccctgttcacccttgctttgcaaaattcttgtag
gatgccagggttgggagtaattgtttttcacaagagtcaaaccacgcttggtttcttgaagaagcaagctttt
ggaggggtgggtggcttgaaatctgttggatctgggtattaggtgatatacctttgacataaaggzcaacactga
tgcaagcagcagctttccttggaaaggcagggagaaagtgaaggccagactgatgagcttacactgacctg
cagacccttctccattcccaggcatgtttgggtggcagagttaccttagttgggggttaggctgttgtctggt
actgtgagagagaaggaagaagaagatatgatattaacaacaacaatacatatttatatttgaaaattaaat
gcactaatcacctatagtgtctattaacaaatttatttgttcagcaaatgtttgttaaacaccaatgtact
gtggaaagtactagggtgctggacagaggtcaaaagactttttaagaatctgccaccattaatgatctctttc
tgcttggcattcaaggctctttgaaataagactgtgaccactttgatagttttgtcctggattataagaca
catgctcgaaggaaactatagctgggttttctcaccagactgattaacatatagtatgggttgggtacctgttaa
atgagctctctctacaggttttctctacttttaagaacccttccctggcattgtttgggtcctcatgtg
ttctggaatctcatgtccatctcatgtatttgccttgggttcacacttgatctctctgctcatattcctcacct
actgaaattttaccatcaaccagaccgtgggtgatggaacacagcatgggctagttcttcttatatatgat
ctcatttaattttcactggaactctgagataggttagcatttcagcccactcaagttgactcaaatttttgta
atcatcaatataattttaaaataactttatatttgacctgtaattggaaaaccaatatgagttatcataaatg
aaaggtaattttaaaataaattaggatgaagacaaattatttttctcacagtctatgtataagataaactat
tgggtcccaaaggccccagctgacaatgagacttctcttactttgttgaaaaggggaattagcaagcattaaa
gaggtgtcaaaaagaagactaaacaaaagcttactccttttttttttgagacgagctcttgcctgttgccc
aggctggagtgagtgccaccacctcggctcactgcaacctcttccctcctgggttcaagcgattctcctgcc
tcagcctccctagtagctgggattacaggtgcatgccaccacacccggctaatttttgtatttttagcagag
atgggggttgcgtcatgttggccaggtgggtctccaactcctgacctcaggtgatccaccacacccggcctccc
aaagtgc tagaatcacaggtgtgagccgccgacccggcgcttactccttaatgtactaagaatgttatat
ataggctgaagaagtgtgaaaagaaccatattttctcatgatgtgggttcaatgtttaatactgtgcttgtt
catctcctaaaatcctctgaatatcacttaaatcatcctgtgtaactctcccacatttggggaaatactga
gcttgcctattattatattagccccatatttcagatgatgcacctgagccgaggagaagttaaataacttgt
tcatgggtccatatttggctaattggcagagccagggattcaaaactcttgtctctctgactcccaggtttgtg
cttttcccacttgggtgaattttctcatgtacctcctccataaacacctttcctagaacttttaaggaatgct
tcccctgttctctcatagcatttttaagttaagttgccaaagtgtcccagtttgcacccctaccagcaatgtg
ggagaaaatagcaacatatttttgatgttgggtctagtgttacgggtttcttctgctctttgggatgtatat
tccatggctactgaatacccaagtcccaaacagttttcttagactcagaaggtcatccacttcagcttcttc
atcaaaaagacatattttctggctgggcacgggtgggtcacacctgtaatcacagcactttgggaggtggagg
caggagaattgcttgaaccaggaggtgaagttgcagtgagccagatcgtgccactgcctccagcctgg
gtgacagagcgaaactctgtgtcaagaaacaaaaaaaggacttgttttctgttccattaccacagtggtta
gaatggcggtgctaaatttattctccagctgccattaaactgcaaattaaaatcttagtctcttgccctcttaa
tccaggcttcttcatactataccagaatttaggataactattacagtgccctttataggagagaaagaagaa
attgtgtctgtagatgtctgttctttcagcttaaaatggacactgaaatgttaaatattggactggcctca
tttattttctcctgtctgttgggtccaatttgaatcttaaggcgtctttcaactggaattttttgtttctctca
actaaaatgttctttgtaagtttgaatcagaacaaaatcctgaatgttgagggtttcctaagggtgtttt
ctttatgcaaaagcctgaaaccgatgttgatgttgggtgcttaaaataactgtgaatcaaggcaggggtt
ttatttttatttttttacttttaattgattgtgttaattatagtgaaaaccttgagttcacggagaagaaa
gcctttgggtcaagttattgtttattaagttgtcagctctgttgcaggatttgcaaattagtggaatttagtgc

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

catttttcagtttacaattccagtcacattttacatgatcagagcatggctttcttctctgtggagcaaata
gagggctgtctgacacttgggtccagtggttccattaagcagagtggatatgtccctggagctctgcagaga
agggcatggcactctgaccccagatggcactccgttttgggacattgtccaattctagttcatagcatatgt
gaccaacaccagctctcacctgatgtaaacacttagcgcggttgttgccttgggggattggattgtgtgaattt
ttcaaaactacagttgacagaaggaggctacaaaaatgaaacccaataattccatttttgggaattattcc
cacttttgttccatttttcccactttgttctttggcacacagaatgtttgatttgtgaaaatcttaataaca
gtagttttttctataaggaacactcagaatcttgataatattggaataatacagatcctttttaggatcct
ctcagacctcatataatagagttcatgttagtcaatatttaagaaaaaacaccttaagtttttgttttccag
aatcacaggttaagtggatttaaaccttgtagcttattccccctttcttcttaatttagtgaggcagccagc
gagaggggttgttttgggtattctaaagaaggagtttgccttgtaagttttggagggcaagacttagactctg
tgtctctgtgcttgccttggaaactttgattaaattgtcactaaccgagttagctggccctcgccgggtgca
gaaatagaagtgtcttgcacacatgacatatgactgtctcaagagctggctgggtgaaaggacgttctggaga
aggctgccgatactgtatgaactagaaactggacaagagcctggagattggataactcagtttggcgcaagt
aaagggaataaaagtgttaagggtggcaaaattgtatccaggtgtttataggctccctgagttcctgacttga
gcctatctatgggttttagagttcaaggctctttaccagtgctgacaatcttatactctaggttgaacctccg
gggaagggtgcccttgccttgatggcatgtttaccaggggttctagagcctcaatcacagattctctctagctc
acatgaagttaatgaaaatgaatgtgcttccctacaaattagagaggctttgaggaaaaatcagattaaatg
cactcctgcttgaacttatgtttcttagaacacagctggaaattttgtcacacaaacctttactttcagtga
catttcttgactgggttgttactgttagtgaatctgctttaactatcttttcttatcgctgaggttttacttc
cattctacatgtgattgtggagcgctgcgtcattgtgggttcagtgtagtggagagtaggaagatggtgaga
cacagtagcttgttgcacattgcttaatttatcagggatcactgatgagtttagtacactagagaagattgta
ggtagagctgaaaagatggaggaattataaggctcagatttctctcttttttttttttttaagatggagtt
tcaactctgttgtccaggtcgaattcaatggcatgatctcggtcactgcaacctctgcctcccggttca
agttagagcttgatgggtctcacttgatgggttctcctggctcagcctcctgagtagctgagattacaggcacca
ctaccatgccagctactttttgtatttttagtgagagatgggggttttatcatgttgaccaggctgggtctcg
aacttctgacctcaggtgatcacctgcctcagcctcccaaagtgacagggttacaggcatgagccactgcgt
ctzggccaaggctcagatttctaatagagatzttctaatggacatagaggctggaggaaatgggatzggaca
zggazaaactgagtzczaggtgccaaazzaactttagggggcggggtgcggtggctcazzcgzcczctgt
aatcccagcactttgggaggtcagggcggtggatcacagaggtcaggagatcgagaccatcctggctaacac
ggcgaaaccccgctctctactaaaaatacaaaaaatttagccggcggtgttggcagggcggtgtcggtagtccc
agctactcgggaggtcagggcaggagaatggcgtaagccgggagggcgagcttgcagtgcagcgagatcgc
gccactgcactccagcctgggagacagagccagactccatctaaataaataaataagataaataaataaata
aaataaaaaactttaggggaggtggcagtgctatggaggataggtgcaacctctgtgagaatgtagagaaa
atagtataagtgagtggtgaggaccccccaagaggggttttatagtaaaacaatggtcagaagtggcaacag
gataccgtataatgctttcacctctaccaatgcactgggtactggagagcgctccagtttgctctggaaagg
ccctttctgtggacaaagaatacagaaaagagattcctttaataaaccaccactctgtgtcccacccttga
tgaatacttcaactgtgaaattgccagaattaatcatggtaatagctactgtacacttactttgttccaggaa
ctggataaatgttttacatacattatcagttctatttttggagaagatacaggggctcagagtccttaggg
ttccagagctgggtgagtagcagagtcaggattcaaaccagctttctctgactctaaaacctcctttcttcc
tggtgaaacaattaactcaagacaacaaggagtttaaggatttggggagtttctgcatggtagaataagacc
caaaggaaaagaaaagacagtgactaagatttgggttctgctgccccaccaaatagctttgagcactttc
ataaatataaatccttcaggttgggagaagggtgaacatctgaagacactgattcttcagagatgtaatcca
aacaagtgatctttgggtgatatggctactaaaccatttatcccaaatctcttggaaaaccgtcctataa
cagcaggggaacattatccagccaagtttttctgcaataaagggttgcctgatagaggcttgctgcttgtgt
ttctzgtzagcztzcagggtgtttatgaattcactaatcccttcccttcagatcccttttattctgggtgtta
tgattgtgactgzaaaaaaattgatttttttctatgacatagaatgttgaaagggtgatttcttttctaga
ggaaagattcttttttctatgtgctacatacccccgaccagggaaaaggcaaatagtggtattgtttgct
gaagtccttcttgaagggttgccttgggtgttgccttagtggaatcagcaggggaagagaggctatctctzaa
cattttgttagazgtttcttctgtagttctatagtgatgaacaaggacttggggtacaggacagatctgct
ttcagaaatcctggctcttgcaggtttagaagccctgagaccatttagctgggtggcaacgggaatgtgag
gggtgataaataggatctttgggtgtccaagtatcagtgacatgatatagatggagttaaacctttaggatct
ccttattttatttgtttgtttatttttgagacagggtcttgcactgtagcccagggttgaggtatgggtggcatg
atcataactcactgcagcctcaaactcctgggtcgaagcgatcctgctgcctcagcctcccaagggtgttagg
attacaggcatgagccgccacaccgggtcaggatctcctgtaaaattatattgttgacaacatgaagaatta
tgcttctcaaaagctagttatagatttgtacaatattcatagatttcttgtttcagttttacaaattcata
gcccttattttgaaaattagctatttagcaataatttgccttaggaaattggatgtgtattcaagtgaagaa
ggaagtacagttacctattatcttattgttaactaacaatcaagtgtgtgatttgggtactttaaaaa

ctgcacccaagttacagattatttgaattaataaaaattcactggatctatatatTTTTTaaacggacagtgtg
 atagcagaacctcttatagaatzgatagaattcctctggaatgattggataacttcatttcaccttgactt
 ttaccttggaggatttcttacctcttgggtctctcaaatTTTgactattaaaatgttgcctTTTaaaatagg
 aacacagtttcaggggggagtagccagcccatgaccttctgcaaggccccctaactcaaggttagttccctg
 gaactgtggtttatggaatgtttcaggagtgtagggagggtataatttaaggctgtcctagcaaggataacct
 taaggatagagggccagtagcatctggaggccagaaaagttazaactgaggcagtcagatttagcttcazgg
 ctcaattaagctgatgggtcagcctgggagaaaattgcaggatgactctcaatatccccctccccccccacag
 cagccacgatctgtctgtctttaatcatgggtgcagtgaacctgttcttccagggtgtcttggccttcagta
 accttgttaggcttgtccctgaacgtggctaccgatccaaagacacatgatcagagagggaattagagaaca
 gaccttTccaaagcaagcatgttctgttgggtctagaagtttcatgtcctaataattataggacctgtgca
 tctctctggagatgaggcacatgagtcatactgtgattcttTgtgtcaacatctcatgaataggca
 atcagagcttTggcaccaatgtattttcagttcatatctgatgtagttaaattccacctcctgtcttTgtagt
 tactggcaagctgtttttgatataagacatctagaacactgtaaatatataacattttttatttTgtctattat
 acctcaattacgaaaaagacatctagaagcaacctcatcaagagagataactgaggccgggcatggtagctca
 cacttTgcaatcccattacttTgggaggctgaggcaggttagatcacttTgagggtcaagagttTgaaaccagcct
 ggccaacatgtTgaaacctgtctctattaaaaatacaaaaaagttagctgggctTgggtgggacactgt
 aatcccagctactccggaggctgaggcaggagaatcactTgaacctgggaggcagaggtTgcagtgagctga
 gatcacaccactgcactccaacctgggcaccagagtgagattacatctaaaaataaaaaaaagtaataaaaa
 aagagagatatTgatagctgtTgtTggaaattTcaacttccatctcacttctTggttaactttTggaagttTg
 ttgaacaaagtTggaatacacgcacatacacacacacatactctctTgtttTgTTaaggtTaatgaaata
 gctgtcatataatcactgtttttgaaagaggagaatttagttgctatctgtacattttTgggtatgtgaactat
 ttggatagaactctgagaaatgcattcagaacaacaaacaaaatcataggagaaatagctaagtgggaaggg
 gcatataagagttgtTgaaaaagtattttctTgagaaaccagctctaattgctaggcaagtcacttTgctttTg
 gggaggcctcagcttctctgtctataagattgcagcaggggtgtagtgggaatgagctctTcaacattTccaag
 agatttttactactaatacagcagtc aaatggagcatgactttTgtggaagcctctcctctTccaccagagg
 ggccaattTctctgtcccagtgagatgtTgacactTgtatgacccctgctTggagactTccctctctTgga
 cctgcccctggctcaggcatgagggtgactgtcaccctTcgataggagccagcactaaagctcctTctTgtTg
 gcagtgttctTgcgggaaggaaaaagaccagccagccattTgttactgcacaagcaacagcttctTggtag
 ctgtacagatacatgcacttcttTcctcactgtgttTccatagacagatttagtgctgtagaagagtagag
 ggcagtcacgggaaggagttcctgtttttctTtTggctatgccaaatggggaaaaatcctcctatctTgtct
 ttttagtgctcatcctctctcccccttttcttcttctTataattctcatctctcctTggaatgtgc
 atgtcaagttTcaaaagggcacaatgtttTggTgaggaagaggtgggagaacacgtgccaggtgctaactagg
 gtcattcattTccccctTcacagccagcttctTgtgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
 gtgtgtgtgtgtatttctTtTgcccagcatcactgaatctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt
 ggaaaagTaaaagtaattttataatcccagctgtcattTaaagccacccczttTgtgggtagcatatggTcca
 ctctctcagttcattTgtcctaaagatgcttcatcagaaaggaataactTccaccccgTtactctctgtcccc
 ttactctgctttatttttctTcgtcaatcctaccaccaccaccactgtttTgaacaaccactattattTgt
 ctgtttcccatccctggtagaataggagccccatgaatgaaggaactTtgcttctgtTgtTcaccactgaat
 ctctaaggtatTggaacacacctggcatgtgataggcactcgataaatattTgtTgtggctcatgggcacct
 gcagagTtaaggctgcagttgtTgtTggaattTataagTggtaatgaatattTatctactattcctctTcca
 agcgatcacacaataatcaggctTtacactTccagttcttaggtctTccaagttatgactTgtgaggtat
 gTtaattatgataatagaaggcagttTattTggTtcagattTattTgatgtgtaattTaccacagtaagactT
 cccctTtaca aaagtatgatgagttTtgacaaatggatacacatgtgtatctaccactgccatgctccttt
 cagTctgtcgtccccctccaccatgaccactggTcaccactgcagtgatttctgtccccctTcattTcacctt
 TccagaatgtcatataaatTggaatcatgcagtatgtagtTttTgtgtctggcttattttTcttagcatta
 ggctTttTgggattcatccaggtTgtcgcagtgaacagtagcttattcctTtttatggctgagtaagtgtccc
 agTtttatttatataattttattTatgaggaggtgtctcactctgtcaccaggtggagtgcggtagcgcgat
 ctcagctcactgcaacctccgcctccaggtTcaagcaattctcctgcctcctgagtagctgggattacagg
 caccacccgcccagcccactaatTtttatatttttagtagagatgggggtTccacatgtTggccaggtga
 tctcaaaactctTgacctcaggtgatccgcccactctggctccaaagtgttaggattacaggcatgagcca
 ctgtgcccagccccagttttattTattcaccagttgatggTctttTcgacaactaattgtTtccagttttTg
 gctattctgtataaggtctTataaatattTcaaaatacctaggatgggatgactgggtcatataatagtac
 TgtataaccttagcagaaactgtcaaaactattTtccaaagtggctctTccattttacaattccacagtgat
 TgagTcccagTgtctccatacacatgctagcactTttaatattTaatTtagTgggtatgtaatgatctca
 ttTgtggTttTaatTtgcatTtctctgcagctaattgatgagTgtTctgcttattTgggaaggtTttaatTta
 gcagTctgtTgtattctgtagataattTaaactTcaaaatatcagTggcattTgcagTtaaaattTcctTaa
 aaaattTggccaaaggtTtccagcagtcactTctgcccagTgccc aaactgtatgaaacaaggctgaggtgtgga

gattgtcacattttggcaaggagtgatccacttgggtgactgatgagaccagagagcgtagcgctcgggct
tgagggtgaggacgggcggaagtgcactgcatggccctgctggccttgggaggtgccagtccttagcta
aagctggcagttatgggaaacagacttagattctattacgttttccaggatgtccaggagtcacctgggaa
gctcagcagtcctttgtgactttcaagcatatggtagaagctgctgaacacagagctccctcttggggata
atltgccc aaatcattta atcaggcttgagaaatgagttaccacaggtccaggagtgctgccacccttgaat
tctgacaccctattttctcctatccgtctcttaattaattaagcagacatccccaagtgcttacgacaagcca
ggacccttttgcatactaaggaaaaacagggatgaaggaaaacagaaatggctctctgctctgactcagaaggta
gaaatcctctttccagccaaagtcttccctagggagcacgttaggaagggtctgaaccacagtgctcagttgca
ggggaggatatacaggaaaggacattgaagaagtggagacctaggtttgagacctaggcattagccaggctag
cagtgcttgaaaaagtgtcttaggacaagagaactcaccagtgaggtccagtggttaggagagcgtgcagca
tattctgagcctgtatacacatctccagggcatttgcttagcaggtggggagtggaagagagtaggctggag
tcacagaaggaggccaggttagaccttggtagcactggactctatgttcaggtgctgaggagctggcaaaa
ggttttaagtcggggagaggcatgttcagatatttgggtctagctgagtaactttgggtgctctgtgacaaat
ggttgggagaccagtgaggtggcagttgcggtcatctaggagcaggatcagagtggcctattgactgggatg
actgtgaagtgggatcccttccagccagtaactggaaatgtgtatgagggcagaagtgagtgactgcattt
gaaacattgagaaatctagtacatagtagtctcttttatatcttttttttttttttttttttgattttgggtt
tgtttgttctactaacttggaaaactgatgtgaaatgtcccttgggttcagttacctgagcagaaggggccc
gggcattgccc aaactctcctcttaggacagaattgtctccagttatgatcattgtgttctgagttgggggag
caaattgtgcaggaggccaggtcagtgccaaggtgggtgggaggaattggagcaggaagcttgccaaagtg
gccagcaaagccacggtagaactttctactgtggctctatgctacttcttagcaaccttctccatgtgctt
cctggagagtccttggagtcagaacctttttcttgaaaccagacactttacttccaagaaaatgctgtcca
agaaaactcatccttcccttcttctcatgaacgttgtgttagaggtgtgtcttctcttccctttagcttttcc
actcagggttttaggggaggtgatattctatatttgggtttgggtctgggtactgcaacactaggctattaag
atltcatccttactgctttgcccctcctatcttccagaaaccacaatggatttgctagaaataatggaac
gtcctgtttggacaggatataaccatttctcagctagaggatattgttggaatgaagaaagataaatgggga
gaagggaactcacattgctttggcacttaaatgaagctactgtgttgggaaattatttatattatctc
gttgaatccacagtagaacacagttgaacaccatacaaggtaagttatgtcatccttattttaccatgagga
aattgatgcttagagagcataaagccttggccaggggacatagttgggaagccggggctaattcatgcctg
ggctcttctgatagttttccttttttaattgtccctcctcattgttaccttggggatttcaagagattca
tgtagcttctaaatcaacgaactgattcctggagagcagcttctgtatgagaaaaatctagctaattattta
tttcagtgctctctggaatgcaagctctgtcctgagccacttagaaaaaacttgggatgacaagcatgtgtc
tcacaatgctgctctgggttggcagtgctgtgctgccagttgtcatcttgaacaaactgatgcagtgctgggt
ttaactcttctccttttttggagtaagaaacttggaggcctgtgtccttctagaagtttgctgagcaaatgg
taaggaaaagaaataggtcctaaggcttgactatttcagagaatttcttgatttattggactgtcaatgaat
gaattggaatacatagtggttaggtgtcttttctctcagacactgcaatttccctccaatctcttgactttt
ctagaagttttaatccaagtccttgttgggtggtzagataaaagggtattgttctactagagactgacctg
gcatggagatctcatttggactcacagatttctagcttagcgttgggtttgtatccatacctcgctactgc
attcttagttccttctgctccttgttccctcatgcccagtgctccaccctacccttggccctactcctctaga
ggccacagtgattcactgagccatttcataagcacagctaggagagttcatggctaccaagtgccagcaggg
ccgaattttcactgtgtgtcctccctccatttttcatcttctgcccctccccagcttttaactttaatat
aactacttgggactattccagcattaaataagggttaactgctggatgggtggctgggatacacagaatgtag
tatcccttgttcacgagaagaccttcttgccctagcatggcaaacagtcctccaaggaggcacctgtgacac
ccaacggagtagggggcggtgtgttcaggtgcaggtggaacaaggccagaagtgatgtgctgacca
tgggagcttgtttgtcggtttcacagttgatgccctgagcctgccatagcagacttgtttctccatgggatg
ctgttttctttccagagacacagcgctagggttgtcctcattacctgagagccagggtgtcggtagcattttc
ttgggtgtttactcacactcatctaaggcacgttgtgtgttttccagattaggaactgctttattgatgggtgc
tttttttttttttttttttgagacagagtcctcgctctgtcgccatgctggagtgtagtggcacaatcttggctc
actgcacctccgcctgccaggttcagcgattctcctgcctcagcctcccaagtagctgggactacaggtgcc
tgccaccatgcccagctaattttttagtatttttagtagagacggggtttcaccgtattggctaggatgggtctc
gatttcttgacctcgatccgcctgcctcggcctcccaaagtgctgggattataggcttgagccaccacgc
ctggccgatgggtgctttttatcatttgaaggactcagttgtataaaccactgaaaattagtatgtaaggaag
ttcagggaatagttataagtcactccaggttgaggcaaaatttacaaatgctgctgactttgtatgtaaggg
gaggcattttcttagaazaagzagaggttaggtctctgggattccagtatgccatttccatcctcagtggtttt
tggccacctgagagaggtctattttcagaaatgcattcttccatccagatgataacatctatagaactaaa
atgattaggaccataaacagtagctcctagcctgctgtcggaacacctcccgagtcctcttcttgggtgaa
cccagaggctgggagctggtagctcatgatccattgagaagcagtcagtgatgcagagctgtgtgttggaggt
ctcagctgagagggctggattagcagtcctcattgggtgtatggcttgcagcaataactgatggctgtttcc

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cctcctgctttatctttcagTTAATGACCAGCCACzGGCGTCCCTGCTGTGAGCTCTGGCCGCTGCCTTCCA
GGGCTCCCGAGCCACACGCTGGGGGTGCTGGCTGAGGGAACATGGCTTGTGGCCTCAGCTGAGGTTGCTGC
TGTGGAAGAACCTCACTTTTCAAGAAGACAAACAGtaagcttgggtttttcagcagcggggggttctctca
ttttttcttttgggttttgagttggggattggaggagggagggagggaggaagctgtgttggttttcacac
agggattgatggaatctggctcttatggacacagracgtgtgtgggtccggatatggcatgtggcttatcatag
agggcagatttgcagccaggtagaaatagtagctttgggtttgtgctactgccaggcagtagttctgatccc
taggacctgggtccgaatcgccccctgagcaccctcttttcttttgcctgcagccctgggagccacctggc
tctccaaaagccccctaatgggccccctgtatttctggaagctgtgggtgaagttagttagtgggccccctctt
agagatcaatactgggtatcttgggtgtcaattctggattcttctcctcaggccctggaggaataataaactga
gactgtttttatttctgcagaggggttctaagccattcacttcccagatgggccaataatgctttgagtaatc
tggagatcatctttaatgcgcaggtgaatggaactcttccacagagggatgtgagggctgtagagcagagt
aactzccctgaaactcagacgtcagctcttgtctctctatctctgaacaccttcccttagagatcccatctc
taggatgcatttctctgtagttagtttctaaagtctcttgttctctgcttttatttttttctctggat
tctaagccagtatcccccttggctgtcttaatgtagcttaacatgtctgtaataaaatgatcatcttct
gagattcaaagggtataaaggacttggagagaatttcattcagtttctcctcaaactagaataatgcttgc
actgtctgtaaaagaacaaaagtgtcaaagcatcttcttgggtcactaaatttcttttttattatagtgtta
cttaaatatttaggaagtzaaagtaggtataaacttctttaggctgttattatataactatatgacctat
acatatttacaaattaagtgcagccaaaattgcaaaatcaataccattcaaattaataccttaaagtgtggtg
aggcagctgttgttcaactgaaaccaaattataagttgcatggcagtaaatgctatcatgctgatcattt
agtttggccagttctatatztatcatgtgctaattgtaattctccaccttttttctactgtatgacctt
aatttgatggcacctgttccatcctcatgagtttgcataattatactgggtgccaacacataaaacaca
aatataaacttgggtcttgaaatcttgtgccagaacttgggtttaaagtaagcatttaaaaaatccatatgt
gtttattagacttgggttagatgactgttgaaatgaaaacaaagtgtttaaactcctcttagagaacttaaa
tataatccctcagcaatatgtatacagatcttcttggagaaaaactgatgtgttcagcctctcatgttac
aaatggggaacctgaattctgaggtctctagtgcagagaacagggactggaatctgtggaztccatctgttt
taataataattgtaaagtataatagataatattatataataaaaataaaagcaaacacttagaatgagcttc
catgtgtgaggcactaactgatttaggcattattaactagatttatttcttttaaggccccgcgatgtactgt
tatttccacatgtttagctggggaacgtgctactcagagaggttaagtaacttgtctgaggtccacaccac
taacaaggagcacaggtagggttcaaactcagataatctgacttggagctggcactctaactcaatgtgcc
taatcgcttttcagtggtgtcattatttgcctattctccatctgagaatattgaagttctgactccttcc
ttgcctttctccctgcctcccggtggttatccccaggtcttgggtgttccagtcctctatgtccgtccttactc
ttattcctttgctacagtgatccagggctcctgcccccttcttattcctggtagagggggccacttgctgg
gaaattgtctccgccatgggttatccatgttgtgtgtccattagttagttagtggaagaatcatatcatgtt
ggcaatgaaaggggggctatgggtctggggtagtctgactgaacctcttattttacggatgagaaagctga
ggtaaaaagcaggggaagggtttcttgagggtcaccagccagcaactgagctgcaaccagaagctgagatcc
ccaggactaggggccagcctcattctgtcccatcacagtgacttttcttccctcctccaaactatttttatt
ttttatttttttgcagctgcttagcagcttgaagttagaagaaaggccagggaaaagggtttccgtgcttag
ccagggaagggaatcctgcaacaggtatgtgggttgggtcattcaaattggggccagactccactgggtctgtt
gcttcttgccttggtattgcagatgggtttaaagtgttaggattagagagataggcaggttttagccaaaggc
agtttgtagccttgtggcagagttcttttaaagaagggaagtggaatgcaacacctgacacaaaggggctt
aagttgttataccactgctgaacctgttttcttaactctcttctgatttctaaagggaagtatatattt
gctgaatcagaaagaaaagtgatttattttaggttgcctgagcttagattgttagagttggaaagatctggc
ttgcatcttgtacagctgacagaactgggggtcaggggggacaggtgccagagttgggtcagtcaggaaag
tagcaccagaaccagttctcctgggtggccctacagttgcagaccttttttgccttgcctctctgtgtatacta
aagcttctatgtctctgaatctcaagttctgactggtagctacttccaatccacctggcttagatttctag
attatattgttttagacgtcagaacctcttaagggttttggggccacttgttagctcacatagtgaagaccag
ccctgcccattaggtagggggaagaagtagcagtcctatgatagctgttgcctgcagcgtatggatgttcatt
gcacagttcctgtctcctgagatcctggagtgtatacgttggcctcagagcccagcacagagcctggccct
tgggacatgcttagtaagtatttactgaatgagtggaagtgtcttaaggcccattagtttgcaggtcttga
ggaggctcccttgcaactaggaagaatagaaagcatacataaagcctgtgtgctgcccaggaagactagaa
acgctatgttcagcctggagctgaatggtataccccagagcaacctgttgaaaggcagtgcttgccctttc
attctgtgtcctgggttgcctggttaactcctgggtccctgctctcctgtacccccattgtgcagactgagg
ggggaccatcagccaggggttagttttccgctgtttctgttaggcaagaataaattgaattgagttgtgaa
gttgggtgcaaagctcagtttgggtccaaagtaacagtttaacttgtgtgggtggcaggtattcagtaaaac
agggtcggggacaggaagggaagagaacttcagagcttccagcatcctcatctgggttttaggctgatccag
aggccaaggtcccatggaacaaactggacaaagtgggggtggccacatggcctcttttcttttgccttttat
tattaattttctcaaatagatctgactagtcagtggtgggtgggaaaatagttaattgtgatttttttttttt

aaactgagtcactctattgcccaggctggagtgagtggtatgatctcagctcgccgcaacctctgcctc
ccgggatcaagcaattgtcatgcctcagcctccgggtagctgggattatgggcacacagcaccacgcctgg
ctaattttttgtatttttagtagagacatggttttagcatggtggccaggctggtcttgaactcctgacctca
agtgatccaccacctcagccttccaatctgctgggattacaggcatgagccactgcaccagccagagtac
cactatttgggcattctttaatgaaaaagaatgaactatccaaaaattaaaactcctcatttatgagctttt
agagaattttacagagtagatggaaactctctgcatcctttcccccacttctagtttcacctgacacatttct
tcctgtccttactcctgggcccggcagcagtggtcatgattccaatcccagcttggccaccatctgcctcag
tggcctaggaaaaactcctttctccagagcttttagttttctcttctacggaatgaagaaagttaaaacaaata
gacattttattgtttcatttggataaatatctattaagcatctattacttgtgggtatgggttagctgggtatat
agtgggtgaagcagctgggcatgagtactgctttcgttagagcttacagttcagtgaggccagcagatgtgaaa
catatcatcacacaaataaaaaataaactatcaactgtgagtgaggattatgaaggaaaaaatccggcaaaact
atgggtactgggtgttagatactagcaggtgtgggttagggatttcatttagattgacaggttgtcacattaaag
ctgagagccctgaagttcaagcaatgggttagccaggcaaaagatcagaggcttagagatagggaatccattc
caggcagagagactgggggtgcctgtcccctaggctcagggaaacagaagaaagccagtgggcactgggtggagt
aataagactggcgggggatgagttggtagtagacatgaccagatmatttagggczcaattctcztggggza
aggagaattztaatttaataattttattttattttattttattttattttattttattttattttttcaa
gacggagctagttctgtcgcccaggctggagtgagtgagcaatctcggtcactgcaacttttgctcc
tggtttcaagcgattctzcctgcctcagzcctcctagtagctgggattacagacgcccaccaccatgccag
ctaaaztttttgtatttttatagagatgcggtttcaccatattggccaggctggtctgaaactcctgacctg
tgatcctcctacctcggcttcccaaagtgtgggattacaggcgtgaccacagtgccccctgagaatttaa
ttttattttatgtgcaagaggattccctgaggtagtcaggccacattgtctggtgactcttgggtagaggg
aacttgaatgacaaaggcccaagaaagcaattgtaatcattacatatatcatggaccattttatgctgttttc
ttctttcatttaacattatttagtggtgctgttcacatatcttaaatcatcttctgatttagaataatgat
ttctgatgtgtaggctgtgttttatagttttgaaagtaataactttgatataccattacttztcttgattctca
cagcaattctgzaggtgtatgcgttgcaattttctgtttcacagatgaagagagtattgttaaatagttaatg
gccgggcatggtgggtcacacctataatccagcatcttgggagaccaaggtgggaggatcacttgaggcca
ggaatttgagaccagcctgggtcaatgtggtgacacccatctctactaaaaatacaaaaaattagccaggctg
gtagcacttgctgtaatcccagctatttgggagggtgaggcaggagaatttgcttgaacctggcagggtgga
agttgcagtgagccaagattgcaccactgtactcctgcctgggtgacagagcgagactctgtctcaaaaaat
aaaaagttgctaagaggagggtggtggtatcttttgggtccaaatctactgtgggatgatgcctttgacattcc
tgatagctgtgcagtaatccattaacacagtttttataagttcaaaazccctgttgccaacatttagattgtt
ccatgtgtgctgttacaataaattactataaagattctatacatttaattcttttattttttgtattatt
tctgtaggccaaaatctgaggaacaggattactaggttgaagggaatggcccttgaagtgtctgatcagat
gtctttccagaggatccaaccaatttaaatagccaccatcaatgcatgagactttgtagttcaggggaaggca
ggcctgggttttaaaaaatcatttccccctctctagcattttttctgatgtgatccttaagatttcactttagttt
tcccagggtctcattggcatgtatgctgttagggatgggtctaaaaatattttttcttcacattcatatcatg
tcatcccagtgattatttaataaataatcacttgattaaatagtgattccttttctagttatttttgggaca
tttattaaaacctggataggtgggtcatgcctgtattcccagcacttgggaggctgaggtggggggatgtg
cttgagactaggagttcaacaccagcctgggcagcatagcaagactccatctctataaaaaataaggaaatta
gtcaggcatgggtzggtagcttgctggagtcagcacttgggaaggctgaggtaggagaattgcttgagtcc
aggtgggtcaaggtgcagtgagctatgaccatactactgtactccagcctgggcaacagagtgaaactctgt
ctgaaaaaaaaaaaaaaaaaaaaaaaaaaaaazzzzzzaaaaaaaaaagatgtgtaggggagcaattttggagtta
tcatttgggtcatttgatatgtagtttttagtttttgggtgctgatagagccagaatgtaccctgaatttgatga
acattctgatatatgggggagctcattgtccccacttacctttttgcctctcagaatatcttttgatatatt
ttatctgttttttccccattgaatgttattaccttatcaagctcaaaaaagtagcctatcgctattttaagt
tcagttgtgttaaatctataaattagcttgggaaatttggatattaaatgaactcatgaagaagcagagttt
agctctccttaattctcatcttccctttattttatctactacagttctgtgggttttcttttatgtaagaagca
catgztttggctaagttaatgcctagggttttttggtttatgtgtccattctcactgtggatagttctctt
ttccccacattatatttaatttaactgggttttcagagactaatagcaatgctattatttaggagaatttacct
tggttctgattaaacttaccatacttgcaaatcatttgcagcttttttagttaactttgtgagttctcttaga
tttacgaccatgccagaaacagaaaggatattttcatctcttcccttctgatgtttattcttcttgtttcct
ttttttatccccatttatattctcaagaatctctcaataactaagaaatagcgacttcatttttcagcgzcg
agtgcattattttggctaccatgattcagaagcctcttgccaaaggcccaattttattctgctagttttctc
tgttctttgtacatggcccttgcgctgccttaaccttgaattaacgtggctaaatctcaagaatttaagagc
accgtgactgtgctcctcaggctaggaggggaaatgggttcacagagtgactggattgtgggtctatgaacttc
ggcagcagcagcaaaagtccaggcatgaataatcaagtggacagtgaaacatctgtagtggtggagatgttgg
cataactatgaatgatgattcaagagtggtttgatgcatattgaataacatgatgataagtagactctg

bioRxiv preprint doi: <https://doi.org/10.1101/123456>; this version posted January 1, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tgctaagccttctatgtgaaatacattttaattctcataataactctagagcagtggttctcgaccggggccg
gttatccccctacccacccccaccctcacccttccaccaggacataacatctggagatatttttgggtgtc
acaatcctgggaatgtatgtgctgatatttagaggttgaggtcagggatgctgctgaacttcgtagaattca
taggagaggtctctcacaacacctatctggcccaaatgtcagtagggctcactatcaagaaaatctgctctag
cagtgctgctcatattatccccatggtgaaatagcaagatgggaagtgcgaagtgggtgcttcggctactctt
ggagcagctttgactttgggtgagaaaagccttttaaaaacaatgtttcttcccatcttcccaccccatgggg
aggtgtgggggttgggtgggttaggcaccaaagcaagatttagaagagttttctgtaggaatttataatggtaa
aggatcaacttcatttccaagctatttatgaggggttatgtttaggaaaagtgcgaagcttagagaaggagg
agaaatctgattttattaatgagtgtagccataatggcatatcctggcagaagtcaacttgggttcttagag
ggaggctattatgaaaagaaataccctggaacattccccgggttgggaagggtgagttctaggttcaatgatg
ggaagaatttttagaggtccaagataaaaagggaagattaaattttgtctctcatgagttctctgggtcagg
tgggtggaactttgcagacagtcctctttaattcactcatacatgctagtcctccagctcagcaagggtttg
agagagcaggtgtctgtatgctctggtaagtgaaggcaagtgcataggaggttgggggtccataatggcga
agagaaggagcccttcagtcagagtggtttgaatcttggctctgccatttgccaatcttggaccattggggc
agtggtattaaactcttgaatctcagcttctctctgtaaaatgtgtataacaagagtaactaatggattgt
ttgatgattaaatgagttaatgtgtataaagcactcacaacctgggtacatagtaagaccttctattattat
tatcatcatcaatttttttaacctcttttctgtatctgcttacactcaccagcttcagctgctccaaatgg
cttgaagattttttgtttgccctttgtgtcagttgccatggggaagatccattctattttttcagtcac
caacatatttttgagcatctgctgccctacaggatcctagatatgggggctgcagagatatccaggaacataa
gccttgatttaattgggtcagatcagtgctcagcaggggtggcaagtgcaggtttcttttaagtggcatatc
ttaaagggtatgtctctzaacatagctttgtgtatggcagcatgatgggtacaaaagcacacacttaagt
tcagtagatctgggttcaaacattgggtgcagtttcttatgggtcgtaactgtttcaaacctcagtttcttca
cttctaaaacggtaatgatacaacctacctcacagggttattatgaattaaatactggagatgagatacaca
aaacgtcttgagztacacagtagctgcccattatgggtgtaagtattataaatctacaagctgtgaattaa
ttttacctctctggatcctgztgatatttcttagaccattccacctagtgggggcatttctacctgagtc
ccggtgggtgtcaaatagaatgtcagtggtggcctcctgagttgggtagaattgggtgctcatctcaaccgcgt
actgactatctctgtgatttacccttctccagccttagccttgctacatataaaatcaagacaataatgtt
tcttatctcacagggttgtcctgaggattaaattaaagtaattaatataaaatgtgccttgtagcatattgggc
cctaaataaacagtagtactatttatccttaaagtacaaatggtagtttcagagcttcaaggctgatggct
atttatcttactcatactctttgttttagcttcatttttttccctaatttcattagwattttcttttctctt
tt
cgatcttgggtcaccocaaacctctgtctcctgggttcaaacagttctcctgcctcagcctcccagtagctg
ggattacagggtcccgcacccatgccagctatttttttgtattttcagtagagatgggggtttcaccctttt
gaccagggtgggtcttgaactcctgacctcatgatcaaccacctcagcctcccaaagtgcgggattacagg
tgtgagccaccacgcccggcctcataagtattttctaaattttatttacagtcagtcattttaaaaggaaagt
tgtattcctgtctttgttaatatattataagtgtttttattcagctacaagcttgggaatggcatataatttg
tattctgcttttttcaacttaatatattacatggctaatgatttctgtgtttcataaacattattctgatgatg
catgatataattgttgagtacatgtaccataattgaatcatttccctattgctatgcaattaaagtgtttcca
atattttgcaattataatgtttcaatgaatgaataactttatgcataatagctttttgatattcctaagttcag
tttcttaggatgaatttccaggaatagtaattgggcaaatgggataaacatgactcttgaatcagttatgtt
aacattgctttcccaaagggtcactgatttatatttccgtgttcatattcttttaaccagctcatttac
tcaccaaacatttttaagccattatcatgtggtaggcttagtaagaagaagtgcacctaaaggagagaagct
tatataataatagggtccctgggtgtaccaagtgcgtgatacagacacaaagtacctggggaaattgagatgag
ggagtcctgggtcagctgggagaaaagttcattttcatagagtcaggttttgttctttggcagaaaagaaaa
ttgctttcttccccacccccacccccagctttattgaggtataattgacaaaataaaaattgtatatctttaa
gatatgcaatgtgatataatgtatatctcaacttaaaaaataagctacagaataaaaagggtgtttgctatt
aaaaaaaagaaaagggtgaatgtcattcccaagcttggaaatttgagtagttgtgctctttgggattattt
acagaaatattagcaagaccagccccatctttgggtcttgagtactccactgtcagcatgctttcttccagag
agggatccatttgctttatttttcttctgttgtgcccgtctatgcaaacatttcttgatagttttatggta
acagtggttttttgttccatgagatzaatttatcatgctcattgtggaaaatttagaaaagacaggaaagt
attaaaazacatcacttttttttttttttttttttttttttttttttaagzagacagagttctgtctgtcgc
caggccggagtgtagtggtgatctcagctcacagcaacctccgcttcccagggttaagtgttctcctgc
ctcagcctcccaagtagctgggagtagcagcaccacgcccgggttaattttgtatttttagtagag
atgggggtttcaccatgttggccagggtggtctcaaactcctgacctcaggtgatccgctgcttggcctcg
caaagttctgggattataggcaggagccactgcgccagccacacctacgttcttatcatctagtagtaca
ctgtcattatcttgtgtatttcttctgcccagtcctcactctgatcatgagtggtggtgatgcagtgga
tctcgggtcactgcaacctaggccttctgggttcogagtgattctcctgccttagcctcctgggttcaagtga

ttctcttgcccttggcctcccaagtagctgggattacaggcatacaccccatgccatctaatttttgtatt
 ttagtagacacagcgcttctactaaaattttgtatttttagtagagatgggggttcaccatgttggccaggc
 tgggtctccaactcctgacctcaggtgatccgcctgccttggcctcacaaagtgattacaggcatgagccact
 gcatccatcgccaaaaagattttttaaagaggttaattgtagaacatatcaaaggctcttggaaataaaaa
 acagttttttaaataatcagaaataaaacaacaataaataaataaaaaacacccaaaacaatctgaa
 gcacgagcacctagcagaaagggttcaattatgatctattcatagagtggaaatatcaagtagacattacagga
 catgttttaagattatattttatgtcatgggaaatgctctccagtgatgatgttaaatgaaaaaacagaata
 caaaagtatatatgctgcatagtctcaatattgtagagaaaaaatatttttatgtatgcatgaaaaagac
 aaaagatgttaacagagatccattgttacttcagtttacttagggattgtctctgggaggttaggattaagggtg
 atttataatttacctttttaaacttttctgtatttttttattttcaaattttccataaaaaatataaggacttg
 aagatcaagaaaaaatttctgcttggctcagtgagctgggtcacgcctgtaatccagcaggttgggagcc
 ctaggggagaggatcacttgaacccaagagtttgacgttccagtgagctatgatctccggatcgtaccgcct
 ggacgatggagcaagaccctgtctcaaaaaaaaatcttggctttttttttttgtttgttttgagacgga
 gtctctctctgttggccagctggagtacagtggcacaatctcagctcacgcgaacctctgcctcctgggtt
 caagcgattctcttgcctcagcctcccaagtagctgggattccatgcaccaccactatgccagctacttt
 tttgtattttcagtagagacaggggttcaccatgttggccaggctgggtctcgaattcctgacctcagctgat
 ccaccggccttggcctcccaagtgctgggattacaggcatgagccactgtgccagcccaatcttttgctt
 tttttaaanaaagaagacaaaaagggttttataaccagattatcttggctgtgtgactctgaagccacagt
 tgtaagttataattactctgaaacacaaggccctgtgactcttttgggctcttgggtgtttatcttgattac
 aacgttggaaatatagaaatgaaaggaatgggagaggtgatagacttcaggcagtgtaactagttgtctgaac
 actactggctcaattatattgtgtctagtgtattccatcttgtccgtctgctaatttatcgctggtaactc
 actgaggcaggggttttcccttggagaaacctcattgttttaaccagtgatcatgcttgtttagaagttcaa
 tgatcttttttaactcatcggaagatgatgaccagacctggacagatggggaaggacttgcactctctct
 ttacagtcctgagtgcacacaggtcaatatggaactatgtgtgaattttcattgtctttgagagccctcttc
 tctgccccataggagagcagcttctgtgtgcaattagaggagcaagggttgtgtgtatttagcacagcaggttg
 gcttggctcctctcctctcaacatagtcaccacatacctggcactatgctaaggctgggaatgcagacagatg
 ggtgcctgcttccagagtgcctcaatgtgtgtaggaagcagcaacagaaacagatgatttcaggagctccag
 gaaaatgtctacaggaggtgtgtgctgggttactggagtagcacaggaggagggttctagctcaggctgag
 attttagtaaggaaattatgccacgatgaatcctgaagaatgaatagaagtgaaccagataaagcacgata
 ggaagcatcttcccttacctaagggaagacacagaggtatatggaatggtatgttaaaagggttgggactcca
 aacagttctgtttaaagcttagagagtggtgggagagactggagaagttgatttaattagtaaatgaagttgtc
 tgtggatttccagatcccagtggtcattggatatccatattttttaaatttacagtgttctatcttattt
 cccactcag**TGTCAGCTGCTGCTGGAAGTGGCCTGCTTATTTATCTTCCTGATCCTGATCTCTGTTCCGG**
CTGAGCTACCCACCTATGAACAACATGAATgtaagtaactgtggatgttgcctgagactcaccaatggcag
 ggaaaatccaggcaattaacgtgggtctaaattggacttttccaaagatgctgtcttgggaaacatcacaca
 tgctttggatcagzaaaacctaggcttctaatttgttgataaggcatgaactcaggagactgtttcagtc
 tagtgaatggtgataattgttaattataacagtagacaacatctcttttacacattttaaatacatgaaatag
 aataaccttactgataattttagaaagtgggtgattaaaagcacatttaagataatgccttaacacctagtct
 tttccatagcatgatgtcttaatacacacattgcaaatcatggaacacagaattttaagcagcatttgtgta
 gaacttctcagtttactaatattatttttatttctcataacaaccttgaatagaactcagatcatctg
 tcaatcatgtattttgataacagcctttacagtgagcatagaaaatacagtagtggttcaacaacacagctc
 cagatgtcaggttatctgggtatgaattctgggtcagcattcactaagcatatgacctggacaagtgtatt
 taagtttcttttaaacagagaatagtaatacctacctcatattattattgtcagtgatcatcttacaatca
 cagttcttctcttagggctgggctcagtggttgattgacactgcagaaatggccagatctaaaggatcaac
 atttacgtagctgggaaatgtagctgggacttcagtttactgcctagtgttttctaccactaagcag
 ctcagttccataccctacgagacccacaagcttatgagatactgttcttccaggaaagcagtggggcccaggg
 ccaccttttaattgtgttttcttggcctgggtcccatcttctcacaatatatagcaacagttatttacttgc
 gattttctaatgcacatcacacatagtcatttaaacacacacacacacacacacacacacacacacccctc
 aagaaacattttctgagacgtgatttctgatttcatcaaaaaagaaaagagcggggccaggcacagtgaggaa
 gtcaagggtgggtggatcacttgaggtcaggagtttgaaccagcctggccaacacggtggaacctcgtctct
 actaaaaatacaaaaattagccaggcgtgggtggcgcacacctgtaatccagctactggggaggtgaggga
 ggagaattgcttcaacctgcgagggtgaggttgagtgagccgagattgcgccattgcactccagcctgggc
 aacagagtgcagactctgtctcaaaaaaaaaaaaaaaaaaaaaaagcataaaactgaaatttatatgcaatttat
 atgctgtgagataattctgttttctcttttggacccccaaagagatttttttgattgatgagcaaatatcat
 ttagattttatttaagcattatgccaagcaccactgaagtataagtttcaagggaactcagttttttca
 tctactagacgaatgattttctggaatgattacaagcaggcaazgatggtgzztagtggaatagcaaatgt
 cttcggcatcagacaagtgggggtttgtttgtatcctgcctctgccttcaccgaggttgtgtatcttgggca

gattgttgagttttaacctagatttctctgactccagatcataaattttcagaaaagttctgaaattcttgt
atatactgatggtaaatgagacttttcttacatctatgcacttcttgtttgtttgttttgagatggctct
gctctgttgcccagactggagtgcagttagtgcattctccgctcactacaatgtctgcctcccaggttccagt
gagcctcctgcctcagcctcccaaatagctgagactacaggcatgtgccaccacgtccggctaatttttgta
tttttagtagagacagggttttgccatgttgaccacactgggtctcgaactcctggcctcaggtgattcgccc
gcctcagcctcccaaagtgtgggattacaggcatgagccaccatgccgggccatatccatgcacttcttgc
aaccttaccttcttttctcatcacccctccagggacctagttggaagagcagagttaaaagttaagggtgaaac
ttggagaggtgtcttgtccctaggaacaaaggactggttgaaattctctgtaaattcttcccagttcaaac
cagagttaacaaggtcttaaaaacttccctgggtcctgagagcccattatattttacttgtcttctctgta
caccacctgcctagtctgactcttcttgttttgcaaataggatggggcacaacgtacaaggaagggcctt
tgccaccctgtctaagggataacctgaaataccttcaccatcactgcctgtgctgcttttcacctatgcc
gtctgtctacagtgccagtgtctcctggcattgaaaggggagaatcttttggctcctttgagtatttgggtgg
gttacataaatctccctgaatgaagagcagctgacttaggcaaggggccttgtttgggttttcttgaactat
taacaggaagataggagattaactgtgtaaatgttcaataggccagagtccctgcagaggggtggccacagt
gatcagatcttatcacatccttgccttgggtgttgcctctctgggtggagataggatagaaaagaaagaaag
accctatattgaaatgcaaagtgcagcaagtctgactttggattaacttctcagccatttgcataaaaat
aaaaagatgaataaaaacaaggttcccaacttggagggaggtggtagctgtgagatggaaggagtgttcttgc
tgggcaacagcagagtaagtgttgggttagattcactccacagtgctggaaaatcctcatagggtcattt
gttgagtcttctgctacaccaggcactctgcaaaaacgcttgcctgcaaggtctcatgcatgctcacca
cagctctgtgaagttaattgtacttttatcaccattttacagatgagaaaactgaggggtatgggggtcaatga
cttgggttaaagtcaactgcttagcaagctgcagggactggatgtgaattccaattgggttgactccaaagcct
gtgaagctacttgttcttccaccacctagagctgtgggtcttctgataactgtgaactcttttgggggtcacaat
agccctgagaatatgatagaagcaggagctctggcctttctgtccatacctgaacaggtccttgggttaaga
gccccctgctccagggcctattaatcttgatcctcataagcagcatccatgtattacggccgcaaaccaact
gtgccagaccgaatcctaggaccaagccccaaatatgtcccatcatccttttggtaagaagctcattgtgaaga
aagaaagaggagagcaagaggtgacctagtgcattggggcctcattgttttaattagtacaaaacaacaat
aataacaacaaaacccccgaagcttcacagatgacatcagccccaagcctgtgtgtttttcaggtgccctt
gaggagctttgtagctggcagaggaggtgaaactgacaaatgtttggcagatggaggagagtaccagagggg
tttgagatgagctaaattccaatctaaccgcaggtgtgaggaagaggcttggattgggaccatggagatggg
gggttctactcccagtcacgccagctgactttgcgagtgttctttgtcagtcactttatcttattttat
ttttatttttttgaatggagtttcgctcttctgctgccagggctggagtgaatggcgcatcttgggtcact
gcaacctccccctcctgagttcaagcgatttctcctgcctcagcctccagagtacctgggattacaggcgct
gccaccaagcccatcgaatttttgtatgcttagtagagacaggggttctgccatgttggccaggggtggcttg
aactcctgacctcaggtgatccgcccaccttggcctcccaaagtgtgggattacaggcgcgagccactgtg
cccagcccacttcatcttaccgtagttacctccttagagtatgaaaaaataggccttagggcatccccagtc
ccctctatgtctgagagctgaggctggctgtcaaagaggaactaaggatgccagggactttctgcttaggac
ccctctcatcacttctccaacgctgggtatcatgaacccccattctacagatgatgtccactagattaagaatg
gcatgtgaggccaagtttccacctgagagtcagttttattcagaagagacaggtctctgggatgtggggaat
gggacggacagacttggcatgaagcattgtataaatggagcctcaaaatcgcttcagggaattaatgtttct
ccctgtgtttttctactcctcgatttcaacag**GCCATTTTCCAAATAAAGCCATGCCCTCTGCAGGAACACT**
TCCTTGGGTTTCAAGGGATTATCTGTAATGCCAACACCCCTGTTTCCGTTACCCGACTCCTGGGGAGGCTCC
CGGAGTTGTTGGAAACTTTAACAATCCATgtaagtatcagatcaggttttcttccaaacttgtcagttaa
tccttttcccttcccttcttctgtcctctggagaattttgaaatggctggatttaagtgaagttgtttttgtaaat
gcttgtgtgatagagtctgcagaatgaggggaagggagaattttggagaatttgggggtatttgggggtatccat
cacctcgagtatttatcatttctgtatgttgtgaacatttcaagtcctgtctgctagctattttggaatata
ctatatgttgttaatgatcatcatgcagcagacgtgcattgaatgggctggctctaggagctagagggtagg
ggctggcacaagatgcattgctggaagggctccttggccataagaagctgacagccaaggctaggggagtctt
gtcttctctgcattcaggtcacctctctcactctgtcactgccccatcagactacaatgtctgcaggtcttt
ctccctgagtgtagctcctgagcaaaagcaggatgctgcccccttccctttgtattcctggctcctggctt
cagtgcttggacataagtatgggcataataagtgctcccccaaatgagacattgaggattcttcaaatgcaca
ggaccgtgatgtgagttaggacggagtaaggacgatgggatgtggctcaggacaatcctgaggaagctgcag
ctgcggcacgcagggccacactgtcatgttcatggaccctagactggcttcttagcctccatggggcccttc
catacacaatatataaaaattatatttcatgactgzcattgggtataaagatgaatataatccagaccagatt
catgattattcatacatttttagtgatttaacttttaattctgcttttaaaataaattaaaacattctaata
tgcccttaagagtatcccagggccagggcactgagcctactgtgggtcatggataagztggccctgggggc
atgtgtgtgcattgcattgtgtgcacatgcattgatgagccgggcttgaaggggtggtaagatttgggtgtgt
agaccaatggagaaaggcatttggggcagtgatgatgggtgggggaggggaacatgggtgatgaatggagctgg

gtgtggggagaccatgggaagtgggtttagggccagcctgtggaggacctgggagaccaggttgagtttctatgcac
ttggcagtcacttctgtaaagcagcagaggcagttggcctagctaaagcctttcgccctttcttgaccctt
tacag**TGTGGCTCGCCTGTTTCTCAGATGCTCGGAGGCTTCTTTTATACAGCCAGAAAGACACCAGCATGAAG**
GACATGCGCAAAGTTCTGAGAACATTACAGCAGATCAAGAAATCCAGCTCAAagtaagtaaaaaccttctctg
catccgtttataaattggaaattgacctgcaccaggggaaagagagtagccaggtgtctggggcttggtccca
ttagatcttccccaaggggtttttctccttgggtggctggcctgtggggccccctctccaggaggcattggtga
agaaactaggggagctgggtgccacagacagtgatgtactaatcttctctgggaagacagaagaaaagtccc
caggggaagaatactacagacttggccttagggacagctaggggtgcatgtgtgccaaactgcatttttct
gaagttggccatatgggtgcagtgaatggatttatagacagagtattctgtgcataaagagcaattacag
ttgtaagttgatatggataagtgaaagtttaagcacttcttctaaaaagagagaattgcaattcattttcccta
atcatttcaattagttctgagtgggcatttgaacttggtgtctttaaaaagtgaaatctttacctctgatctg
gtaagtatccaggcaatttctgtgtgccaccaggaggtatctggggagtgggcattttctgactgaggca
ttggctgccatagcatcagagcagccttccaggcagtgggcctggcaaggggacagaggctgggtgggagcagc
tggctgagtgagccagtaatggcatgtgcatggctctgtagagaatgtagaagcaataatgaagccgataaa
agctggctctgcattttattattatcatgcgcgggtgggttctaacaatgtcagtgataaattactcctccc
atcatggaccaatggctgccactgctccagggaagtgtcttttatccgtttgggtgtttagggagggatgga
gttggctggcctttgctgaaaggcctaccagtttggtttctattttggcaaaagaagaaatgataaagttyta
gagtttaaccagactcagatttgagtttttttttttttttttttttttttttttaaggctacagaactgtg
ctttccttgggcagtaaaagaggcaatgggcaatgtgggacctgatzgacaaazagggaaazzzgzzzzczzaz
zzazzzazzgctgtcttaggggtggcatggaggaggtgctgcttcacagcagagagaggtatggctgtgctt
ggagtgtccactagacaactcctggctgtgcagccagggccactcgagatgtgttctctgacctgcaggtc
ctggctcttcacatgtagtttcttctgggtgcaggagacagaaggtagcaacaaccctgatcaaagcctc
agtcttctcttatttactggagagccccctgctgattgaccagagggcacagctggggatatttctttacctc
tgtagcaagagacagcgtgggtgcagaggaaagtgttagcatacattacctgtggctgcatgactttgtgaa
taggttagtttagcaccctttcagccacttcttcttacctgttaatgagataaaacatgtaattgcttaaaaa
cagtatttggcacataggaagcacttagtgaatatgaattatgattttttttggagtggtgacatctcaacc
aagccatttaaccctcagccttzaacttctcactataaaatagcagctaactgaaatgtaaaactataa
aacctaattgtagtatctggcacatagtagattcccaataaatgagagccagttattcttcttaagacagtgat
gcatttctgagcacctggccttggttcttctgccttgcaatttatgcagcagttgaaatagactggctgatgg
gggtaagttgtcaagcagactttctgatcttagtgaggagactgccttaaaacaacactaatttcttttt
cttttcttttcttttcttttaagacagagcctcgctctgtcaccaggetggagtgcagtgggcagctcttg
ctcactgcagcctctgcctctgggttcaagcgatttctcctgcctcagcctcctgagtagctgggactatag
gcatgtccaccatggcccaactaattttctgatttttagtagagatgagatttcaccatgttggccaggatg
gtctcgatctcctaacctcgatgactgcgccgcttaggccttccaaagtgtgggattacagggcatgagccac
catgcctggccttctttgagaagctggagacatgagttaagtgggtgaagaagccaaatctgtatctaaaaac
cctacagtagtgtgcagagctctgaggagagaagggtcccttagattttgagtgtattattatgtcagtgctt
gttttacatctctctgttcacgcagtatgtccccctttctgccttgcagctgtttcttaaatctttcttct
tttgcttgtcttgcagcacaaaacaggcttcagtataggggggaaatgcacagaaacactgccttttcttaca
ggaaatcagtaactttttactgattttgtttttatttacttattttattttgtttaaatttatttttagtttt
tttttttttagagacaggggtctctttctgttaccagggttagagtgagtggtgcccttagagctcactga
gctcactgcagcctcgaattcctgggtcaagtgatectcctgccttagcctcccgaagggtgggtagaca
ggcatgagccactgcacctggccaacttttctgtgattgcgaatagcactcttgtcaatttcggagagaagc
tgagactggcatatgtcagtatggatccccacttagagacctgtgtttatctgcactgacaccccatcacag
catgatgagcttgccctcctgtgctgtctctccagggtgggaggaactctgaaagctgatctggtttggga
gctttgtcctcattcacctcctttaccacacaccacacttccagggcggggatctaccactcactaaagt
cccattctgggtgttgacagctctaattgttagaaaaatattcaccacctgttatgctttctagagaacaag
tctaattctgttttcttgaatatgtcgaagacagctctcatgtttttcttccctgttttccaaagtcca
tgatttttttaggcaaaatggcctcctttctcttatggaatgttttctcctccatttctgcctctcctctg
gttgtgtttcagtatgtctgtgtgcttcttgaagtttactggaaattatgaaagtattctggcacagaggag
gaagggtttttgcctcccttggtctgagtgctacatttccgttaatgcagctctgagattgtattaggcatt
ttggcattcacgtcaccttggtgactcatattccatgtgcactcaacaaaaattgtgattattttaaatagg
cagaattgcaagttacgtgttctccatttctttgttgattgtttggctttttgaaactaaagggaaaaatgtc
tttttctgttttacatgttttagattccctatgctatcctatcctcccaaaaccatttttagattctgatttt
gccatgtattatatctgatactcctttctcgtcatctagagatgtgataaacaactctctttggcctcattc
cagtcattcgataactgtgggacaaaagactggaagctgggacgtcagtgaggagactaacccacctgtaga
cccttttttctcactataaaaatagcagctaactgaaatgtaaaactataaaacctaatgtagtagtctggc
acatagtagattcccaataaatgaagacccagttattcttctaagacagtgatgcatttctgagcacctggcct

tgttcttctgccttgcaatttatgcagcagttgaaatagactggctgatgggggtaagttgtcaagcagact
 ttctgatcttagtggaggagactgccttaaaacaccactaatttccttttcttttcttttcttttctttt
 tttttaagacagagcctcgctctgtcaccaggctggagtgcagtgggcgcaatcttggctcactgcagcct
 ctgcctcctgggttcaagtgattcttgattctgttagacactaccactcaggcctatatattgtaatcagtgtg
 ggccactgggctcctgcttctgtgatccagttgggaagtttatcttgttcttcccttcagcttgcactgct
 aaattcgctggactatacacaggtgatttgttagatatggggatctctactcaaatactctcatgatttcctt
 ggctagagcatcattttatttccacttattggaagagaccttagagaccagttagttcatttatagataaat
 tagttgattctgtcattcaacccttatatatattgagcgtctcctatatgcgaatcactgttctaagtgcgag
 acacagaggtgtccaaaacaaatatggccctctccatatggaatttctattctagagaagaatctgaccca
 gaagggggaagtgactgtcccaagtctacacacacagaaggatattctgggaataaatcacggctaaac
 cccctgctgctccaggcagttctcctctacagtgcctattgtgctgttttaataatcttcaactggga
 agaactcccatttcaggaattaaagcgtggacaaatctttaaattatccttgaaatcatcctaataagaaat
 ccaaggaggaagtcttacaggggtgcctcaccacttttctcatcactggaacttttagacattttattatt
 ttcttcttaaacagagtacaggcacacaagttgagtggtgtggtggctaaattaatgattgttgaaggc
 agtgtgagaagcattcattcatcttaaatacctatggtgactgcaactcagatgtaaaaattggataaatcc
 tcagaaaccctagggaagtgacatgtctgtattttgtctctgtgagatacagactggcagagataagtgtt
 tctctgggtgagttttgtgtggtatctgggatgatttttaggcagtagctggatgagaacttttaattttaac
 ccacatccaattgcaatttcatggaaattattgcttaggaggatgttcaacaggaaaaatataattaaagtt
 aattcaaaagaaacattttctgtgaatatggtaaaacttgtgagagtagtttgtaaattgattgaagattgga
 aaacattgggtataagagttagtggtgggttttgtattaagattcattttgggaagaaatccatgctgcaccc
 ctcatgaagtgtgaactttgggcatgtgttgattcttctggcccagagtttacctgaagattagctgcctt
 gagggctcactgagcattaaattagatgatgtctgtggatgactgatagtgaaagctcatagcccacagttgac
 acataataaattcgagttgctttgcttcccttctgttccctggctgactgtttggcctttggcacttgttcg
 gctctctgggcttaagtttctttgcttgaatttgaactttctttgggtgaaacaaccagaaaaatgcttc
 agcccagaaacttggtcagtagcttggatgggggatcacctgggactaccaagagatgtgggctgtctgctag
 actatagcccttgaggggaaggtgggcgcttgcctgattgggtccatcctaagccccagcacagtagtgggt
 gcatgggtgagccatttagtgaatcttgtggaatgaaggtgggagaaaaataaaatacctgtacttcacaggg
 attgtgaggggtcaagtaaaagtgtttaaaaaattgtattatatagtttattcccttgtgttagcccaggtc
 aacagagcctacgaataataatgatgacagaagttcttcaaaaagtcttggccttcttctttcacaaaatt
 gccccccagagctttctggaagggcagccatgaaccacagaggcctaaagttagatttactgggaagctaaaaa
 tatttactttatttttcatagctcctttcaaggtcctctctgggggtcttagcaatatgtttacacagtggt
 atgtttttgtaaggtttgcaaaagtaagattttttaaaaatactatcttgttttaaaaagagagccccctac
 caacttgtgtcagcctcagggccccacctgcactgtctcctgccagggcatggtggggcaagaagcactgct
 ccccttccaaagcttccctccttgcctggagtcactcctcactccccactccaagccacctgccatcgctgt
 gccccctctctggtgaatctggcattcttaggtgggctgagaagcagactggcccaagctaaggcctttctg
 atgggggttgttgcctgctgagaatcatgactgggtgggagaaggaggtgaccctttzgctgtcttatttttac
 tgtgtatttccctttcagctacttaaaatgtattgcttagtgatacctaattgggttcattagcctgcttcc
 actgaacatttccgctcaggcatccacttgggtcccaaggcctgctcctctcccatattctgaaatctggact
 acagactctcattcaactccaggttgcactgtggacacagtcctcctcttgagcaggtacctccttgcagtgg
 ttgggagctcctacttggctcatagtgggaagtgcatagtctggagctgaagcctcttgccttccgggat
 agggcgctcctcacatccccctctgagaagttccccagcttccctctgttccccgtttccacacttagcgaggc
 tcttgtccactgctacatcccccatagccagtcctcctcagccttgccattgcttatgctgggtctggaacaat
 tcctagacttgtggggcatctggggaagttctccatcttttttttagctggcatgaccaagtgggtgtgggca
 gggctgtggattctatggtgtgggtggaagccaggtagcctctctctactgtacatggaactcagcaacttc
 tgagtcaagcaagatcttagctctgcaggtgtcttgcctgtccaaagttatggccacaccagtagctttta
 actctagaagcccagtaagtgtttgtgggaccgcaaagatcattttctagacctgctgaatatgcctagaa
 cgggtagggtatggctttcacgctgttccctagggtgacaagtcacacgtttctgggggtacatacacaccgc
 ggtccctgtgaatggcactctccatgagaactgtgatgatttgagttgaatagtgcacagcctacatggttc
 tctgccatggcctggagttccttatcttgccttctccagtgaggactagggtgcaactggcctactttggc
 tctgacttgggggattctgaaataccttttttttaagggttgtggagctctctgaagcttataaggattt
 tgccaggaaaagataagaaatatcttgggcattttgtcactgtgctggagatgaaccctttggaggacata
 tcaccttgttgaggtcaaggggaggaaaggacaggactggcagagagatccggggcagcagcctgccatcc
 cgactgagtatggagtttctctctcccttcagctgcacttttgtgtggagtcagtggtcagctgccacttc
 ccttatgttcatggcatgaatctggcttgttaggcctttcttttttttttttttttttttttttttttttt
 ctctgttggccaggatggagtgcatggagtgtagtggtgtgatctcggtcactgcgaccacccgctcctgg
 attcaagcgatttctcctgcctcagcctcctgagaagctgggattacaggcgcatgccacaacaccctgctaa
 tttttaatttttattttatttttttttttcaaggcagaagaatttttcttagtacagaacaaaatggaatc

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tccatgtctacttctttctacacagacacagcaacaatctgatttctctatcttttccccacatttcccc
ttttctatttcgacaaaactgccattgtcatcatggccattctcaatgagctgctgggtacacctcccagat
ggggcgggcgggcgggtagaggggctcctcacttcccagaagggcgggcgggcgagggcgccccccacctcc
cgagcgggcgggcggtggcgggcgggggctgccccccacctccctcccgagcggggtggctggcgggcgggg
gctggccccccacctccctcccgagcgggcgggcggtggctgacctgctaatzttttgtattttcagtagagatg
gggttttaccatgttggtcaggctgggtctcgaactczczztzgacctgttgatccacctgzzcctcazgc
zzctcccaaagtgttgagatztacaggcatgagccactgcgcccggctctgttzttttggttgtttgtttg
ttttgtttttgttttgagacggagtcttactctgttgcccaggctggagtgagtggagcggcacgatcttg
gttactgcaacctctgcctcccgzgggttczaaacaattcttctgcztzcagcctcccgaaatagctggga
ttacazggcacttaccacczaggcctggctaattttttagtatttttagtagzzazgacggggttttgczc
gttggccatttgaactcctgacctcttztgaactcctgacctgtttgczcgtgttgaactcctgacttca
gzgtgcgttggccztcccaaagtgtgggattacagggtgtgagccacctgcccagcctgtttggcctttct
gatatzgzgczctctgactaatcttttggaaattzagzzztcccagggttatactggattttacttaggg
aaaagggtcatgcztctctggctgtcagzatttactgatagtaactaaggzactcagztgggggtgzzgzac
ctttgattczztzgggtttgatttttggaaatzcaaaaagacgtgagctccaggggagcagggtggctttgggtg
azcatggcaagatagttggctgtggczagggagttagagggaagtgggtagaaaattaacatcttgtazaata
ttzzczcctgggaaatataccttctgtttaagagaacagacttggcagzcczzazzzgzazzztzzzggcc
taggttcaaactcttggcttgatgcttatcagctgtgtaaccttggataattccatacatctctgtgcctcag
tttctcaaagtgaataacaatagtaacctccctcaggactattgtggcaaattaatggacgaataaggggaa
gcacttagtacagtgcctggcccagcataggtaccaggcttgttcttaagctcactgcatttttacaatcat
cataaaatgcaggggatacacacatgaaggagccgaagttcagagaggccaagtaacttgcctaagaaagca
cagctggcaaggggcagtaataggaccagaattccagtttctcgtgctctgttgttgttatatcctaagag
agcagctctgagtagccagaagcttccctaaagtacaggacatggggcatgggctggctgggatgagaaag
gagacaagagggcttctgaaagaaatgccagattcactccacttctggcttcaggcaccgatggaatgttt
cccaaggcccatctagaaagaacatcctgtgactcacagccacttctcattttctgtctgaacccctcacc
cttaccaggcagctgctaaagttgaggttacagcctcagactatattttctgtccttgggaacccagtggtg
tcatctgttgggagatctgggtgatacatgtgtcaacattatgtcatcaaaatggaaattctttgaaatctt
taggtgattgcaattcacgttctgtatgtatgcacttgtcaaaagttttgatttgaggccttagaattttat
atttgaaacctttccactaccatgagttttccagacctgtcaaagccaggctgcatctcagaaaccagggt
ctttgatttccatccagggaagggcctggggccagctgggctgtaagcagggtgggggtggggagcaacgct
gcactgcaatgttgaaatattacttgaactaaatcaaatacaagatcagctttactcagacaagaatagaaa
acacaattgcattcgattacagaatagtgtgtatccccacaaatatacagactgcctttaaaaagtttgaa
tgttaacatcaagaacagtggttgcgtgtctcctgcttttccagcataagggtttatttattctgtgggtggca
aagagcaatttgggagtcagtttgtttctcattgaaagctttccatttctggctcctcttgtcactgttgc
attgaggcacccaaaaggcaatctcagtgcgacactattcaacagactaagttgcaccggataatgataccat
tttacatttttcatatattattacattgaaggcttcaaacagcactagcaggggcaaatttggtattattatc
tccatttcatatgatgaggaaaccgaggtcgaaagggtaatgtctgttgcctcaagattaaagagtaaagtgtg
tacgttgaatcggggtctgactcctaggcttagcattttctccccacactatgctgccatgttgcctattcc
aacattaggaagcataggtgccatccccagcttttgaggccaatatcacgatgaagcatttttaaaacatct
cattaaattgctgatatagtggaaagaaccaagctttgcagtcagagctgtttgggttcaaatttaccact
cgtttctttataggctgttcttaagattctggaaataatgctaataagtgtgcctaagcttggtaaatat
gagtcacttttctgtgcccacaaagcactactatgtcccttaataaattttgttaattttaaaagttagaaa
aaaattaaactattttatacattgtgtatgttaattcttccctagaccagccttaggaagaatctcatcccca
acttgtaaactcatctttttccgttcttttgtgcctggacttctcagggccctgcaggctgattctagtcc
catgttgtgtgggtgtttgaagtgtctgggtccctttttcagtgagagaccagctcatccttgggaactgaat
gcctcaaactctctttttcttctcttcccttctgtttgatgtagtcttctctgtttctggactctgt
ttcttcatacttccatctcttacccttcttttactccttttgtcttccagctgtcctctctcatttttctg
cctctctgggtcttcaggtagagttttcatctcagctatcttcttgccttttctgatgttgggtctttgtttc
ttctctcattctgttcagggtccaaaattcatttgggtcaatgttatgtcttagtggttatcttccatttcc
tctgagcttcaaagccaggctgactgtgccttccacgcctggccagtggtgaccaggacatcctttccctc
tggggctgcactggctcttttgggggaattgttaccattcagggtcttcaaccctcattctagggacttccag
taacttctcccaccttccctcttctcagtaagacatgggtattgtcttattctgtttctgctgctagaagaa
aattctcgagactgagtaattttatacacaatagaaatttacttctcacagttctggaggctgggaaatccaa
aatcaagggtgttggcagggttgggtgtctgggtgagggctgctctgcttctaagatgggtgccttgttgc
atccttaggaggggacaaacgccatttctcactggcagaggactgaagcctctccctcaagccctttt
ataggggtcctcatgtcaggcctctaagcccaagccatcgcatccctgtgacttgcacatatacg

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cccagatggcctgaagtaactgaagaatcacaaaagaagtgaaggccctgcctgccttaactgatgacg
ttccaccattgtgatttgttcctgccccacctaactgagtgattaaccctgtgaatttcttctcctggct
cagaagctccccactgagcaccttgtgacccctgcccctgcccaccagagaacaaccccttggactgta
attttccattaccttcccaaatcctataaaacggccccacccctatctcccttggctgactctcttttgcga
ctcagcccactgagccagggtgaaaaaacagctttattgctcacacaaagcctgtttgggtgggtctcttca
cacagacgcacatgaaactcatgggtatctgtgagggggccctaattgctattgggtgagggcactgctcacataa
cttaatcaccccttaaaggccctgcctcttaatactgtcacattgggtgatgaagtttcaatgtatgaacttt
gaggggggacacgttcaaatcatagcaggtgtgctattgccttactagcaaagtaactctgggggaagatgag
taatgtctcgttccacccctgattccagtgctgactgattctggagtgccagagaggggtggaggctcacgc
ctctgctcacccgacccctctggccatctcctcttagaatgcaaggccagggatttgggttacacagcgcctct
tttgttataggctagctcctgccttcaaggagcactgagaaaaatattatcccttgaaccacaacactagt
attttgggtactgtagccattcagaagtttgttgaactaatacaagtttatattttgtaaaatactagaa
ggaatgtgtggttcttaaagttatgggtattcccattgggtgaaagagaagazctctggcttccaatgcctt
ggaaggttcagggagcataagcctaaatatcctgggttctcttgggaattcttcacagcttgtcatctatga
ctttatccacatttattttgaacatgtttaattttcagcttgtattttctcaaggagtagtgtgtgagtcagc
attcagagtacttaacacacttagtgctcattagtgataacttaaccacacgtagtaggaagctagacctcca
tttacagaaaaatggaggctcagagagcctaactgactttcccaaggccacataactggaaaattctggaac
tgagttttttcagggtatgtcttgggtatccctcgggtgagccttctctgtagagaaggaagtgctcctgggtt
ggaaagcccaggccgggatgaacagcatgggtgtccttaggctgtgtgtaccacacacccctggcctactct
cccctgctctaaggagtcaactgaaggagcagctgcatacccccgcctcaccccttctcctcagcacccaac
aaccaggtcttcccagagagtgctcagctattgggtgaagcaggcatcgtctattcttcaagaagacagcaga
gcatcaggaaaacaaaacaccaaataataccataaagacactggagtgatttgttaccagcccttgatata
acagagaagtatttcaaggctctatttgttaaggctggacaagctcaggctgattaaaaccacaggttagaat
ttgtctcccgctcccaggccctgctcagttgacagctcctagcaggctgggaggcccccaccccaccccacc
ccaggggtccattgaagaagaccatctgggctgggtgagtgattttccctacttctcctttttttt
catttgcctcctccacccacttctaaagcagctgattgctctggttcccttccatgtgagcaaaagagaag
acagatcttgaacactgagctggggactggagggcactgaccgagagtaaccccgctcgggtcagctcagct
cagcaagggtcttccctcctccgcctcctgagctgctctcctcctccttctcctccttcttcccagcc
tcccccaagctccttccccctctcttccctcctcctcctcctccacctccttgcctctatgctcagag
aactgaaggaaactgaatggcctggaaagcaggacctcgccccaccccgctcataccctcgcctgca
cgtgcaaagtgtgttccctcagcttgggtcccagaggcctggacctgggtcccagaggtcccagctgggttccccg
gcttgggtcccagaggcctggatgtgtggcaagaagctgaggggtctgtctcgttcttccagcccacaaaca
cacatgtcacctgttttttgttcttttcccttcccttggccccactttacagaagctcttggctaattggta
ggctcactgttctcctccaggcaaaaccaaggaggtgagctcagcaggtcaagtttttgggtctgtattg
aagagactgtgggtgggagagccctcttgaccttctgctggatagtgactttctcttctgctggcagcagtag
cccagctccgcacctctgtgacagaggtcataccttccagttaccgagcgaacacttggagcaagttagctt
tttccctgaaacagcagcttccagttcttgatcttgggttgcacaacaacctgtgaatgttattagcccc
gttttttttgtttgtttgttttttgagacagagtgctcgtcttgttggccaggctggattgcaatggcgtga
tctcggctcactgcaacctctgcctccaagcaatttctcctgcctcagcttccctgagtagctgggttacagg
catgcgccaccacccggctaattttgtatttttagtagagacgggttcttcatgttgggtcaggctgggt
ctcgaacttccgacctcagttgatccacccgcttggcctcccaaagtgtgtattagccccattttatatc
caaggaaaactgaggttagggagatttaaataggggggaaagtatagcttcatgattacacagtcaggaaat
ggcagagctgactcttgaattcagacccctgtgatgctgtatccagtgctgattttgttttgcctgtgttcc
tttcagtgctcaggttaagaagacctcagagatcatcttctccaaccacacattttactgattataaaatcca
ggctccgcattaaactcattggccaaagttaagtggctaaatgactctcctgttcaggaagtattccctgtcctt
gtgtggaatgattgggtcttctcgtccaggctcatcttccacctatatccatccttgcagtgactgggtggaatg
atcattcctgggtgattccctgtaaaataactagattttatttattctttcataaagttatgagtagacaaatca
gtcagcttcaacctttcctttaccttatagatagtatcagcagaaagaactcaataattctgtcttaggaac
tcagccaggctaatatgtattataatacacacatacatatgcacacacatacacacatgtgcataccttaga
tcacaaattcaacatctcaacttagccgccaagccacztctgtgtcttgtctattctgaatatgctttaa
gggttaaattagccatgaatctgttcagatcagtcattcctgaaaatgatgatgatttagccatcctttacc
gggcaccttactcagtgcccggtattatactaggctctgcagggtacatttgcctcatttaacacctagctcag
ccaatgaggtggatgccattgtcatctcattttatcatttaatacaactcagccttccctctcctgttccc
cactctgcttctagctgccactcactgactctcctcttgatctgggctgacgtgggtgagtggggttagca
gaactcttgtctcctcttacttgggttgggttcaagttcatgtcctgcttcatcttttccctaactcactgc
aaggtgaattttctctctgcaatcggaattcccaaatacatcaggtcattaatttattcactcatttg
tattcattcattcatttagttcatccatccatctattcattcattcattccaatttactgagtgccctcctctgt

gccaggcactgctggcttttgaattaaaaaccctggccctcaaggaatttctattcttctgggagcaggata
tttcatggttgactgczagcttgatctgtgtgtgactagtggaaaacataggatgttgggcacaaaagagt
cacaaaacgccccgccccggtggctcatgcctctaattcctagcactttgggaggccaaggcgagtggatcacc
tgaggtcaggagttcaagaccagcctggccaacatgatgaaaccccaactctactaaaaatacaaaaaggaa
taatagctgggctgatggcgggtgacctgtaatcccaacaactcgggaggctgaggcaggagaattgcttga
acctggggctttggaggttgcaatgagccaagagcgggccacttactccagcctgggtgaaaagggtgaaac
tctgtattaaaaaaaaaaaaagtcacaaaaagggtggccacctaacccagcttgaggatgggaaggccag
ggcaggcttcttgaggaggtgactcatgggccatgtcatgaggatggggtaagaggagaaggtatgtttca
gacatctggcttcttaataattcttggtttccaccattttcctgccagcaacataggggagaagactgag
accagcagatacaaaagccactgtactctgtcctcaccctcctttttctcccttctactaccagatc
tgaggttttgagaaatctcttcttaattatctgcttctgttcatctattgttgagaatctactatgtgcata
gcattatgaccacaaaatgtcctggctgcttgcctcatgagaggtcttacccttctctccagccactct
agggatttttgactctgtccttttccagaacttggctccagctctggttgctcgccatgaagcacttacagat
aaacctcatcttgggccagtgtctccatttactgtctccttttggcttgcttatccttcttctgccttctt
gaattgatttgcttcttctctcccttccagccttgaaagttcctcgaaggcagggaactgtgtcccatct
tttctataaatggcattgttgtacatttggtaaagttagtgaaaaaatcttttgataatcacaaattatta
ccatatattgagcacctaccatgatgagtatctctgatcttcatatctttgaaatgtggatactattacctt
cacttaatttggaagaaaatttgtgctcaggagggtaaagtcaactttctcacagtcacacagctatttcaca
ggagagtgcagtatcaaatttaggttttctggtcccaaaccctgatgtttccccacatcattgtttctca
gacttggctgcaatcacctgaaaaattttattaaattcttttttttttttttttgagacagagtctcgct
ctgtcaccaggtggagtgaatgggtgccatctctgctcagtgcaaccactgctcccagcttcaagtgtat
cctcctgcttcagcctcccaagtagctgggattacaggcacccaccacacactcagctaattttgatatgt
ttagtagagactgggtttcaccaggctgatctcaaactcctgacctcaagtgtatgcccacctcaccagc
caggttttattaaattctgatgcctgggtcttaccctagcagttctcattagttgctccatgggtggtgctg
ggctttgggatttttgtgaagccccagctgagtcacgggcagtcagtttgagaaccagtatactacac
actgttgccttctgttttccacacaggttgggtgctagtgtgtgagtcacagcctacatgggtggatatccac
gtgatcaggctgcaagtcttctgtagtggaggtcttggccctcctgctgcttccccaccagcaaacct
accacaccatgcacgtgcacagccaagggttgggtgactgttaatcagtcctctgggttcttatgtctcatcc
gggggattcactgataggaggctggagcttgatcacatccagatgttctcatcttcagagcctgttaacata
acagaagtcttataaacatgctgagcacactcactgggtctggagagtgttgcaggatgacctaggccattg
caggggtggcctgggtcgccctccacacttctgccacctgcttgggaaggaggaggagctgttttgaagttc
cctggtgcaatatgatgtattgctctggctctgctcaggaggaagctatggccacctagtcagtgagggtt
agctaatacgtgtattctgtttctgttgagggttccacagaggccttttttgaccttctattttatagataag
gcagtggaggccacaactacatgaaatgactcgacaaataaatggatttagaaccaggtttctgactccag
ggtggtgctttttccatgggtgtacagtgattaatgtctaccttttcacaccagtcctcaactgaagacacc
agcttacaccttcttctgtttctcccaagaacagaaagtgaccccggtatgtcgctttctgttctctgggaa
ggcagttccagtggttagaagtcctgttactcctgggtgtggcctggggatggctcctgacatccctgggc
tcttctggacctggccagctaaaaggaaatctcctatgatggtactcagatacttttggaaaccttgtcagc
cctaataccatctcctagtgttttagtattcagctacccttactgggcagtaattctgtgccaggcctgatc
tgggcattgggtggtactaagaacgcataattcctatcctataggcatagtcgtgtagagacaacatgcaagt
aaaacaatgttcttcccttaactgtggttccacacctcctcccccaacattaaaagtgaagggtatgcttattc
aaatgtagattttgtaggctctgcactctagaccactatttcagaatctctggggagctgggccaagaaact
gcattttcgcatgctccctaaatgaagcttaggtgctctgaggtttgacaactgcagtagagagcctaagtc
taacagtgtagagtcacatgtgatgggaagcatcaggtaggttagcagtttgcaggagcactgattctgaggg
acactaactgggcctaagaacagctactggctgtcatgaggaataactaggagctagccatagaggggtagc
agtgaatcatttctctagcgtatgtaaatcttgcctcaatttattctgtctatataactcaatattactgaagt
ttgcctaaagcagaatacacctggatcatacagcatttatgagagactggctgggtgtcaggccctcctgt
tactttatctctgcatgtgacctcttasstccgaggattaaactcctgtcctcattaagcctcacactgtag
ccccattttcagatcaaacctgtttctcttctggtaaatgatttcagtttgcaaagtttgccctctagaggt
tgcttagtgctggccatgtgggtcagttcatgtggtcctgatgagctgggttttatctttattacaaagaa
gttaggtgttaggagagtgggttgggaaggagaagaggttagacagccaaatgagatgagtcagggaacta
tactgtttcgaggtcatagggtcctaccaagcatctggtcagaaacctctcattttggagatcaagaaatt
gaggttcagaaagatgacatgaggtacgcaggagcgcaccagacacagcctccaactctagaaactaaaat
tctggattcttagtgctgctttttctgttttgttgactggattgaagccttttctaactgtactcagaggg
cctattatttagggagattccgtatgaaatcctttagcaatcaaatcatttaaatagggtgatggtttaaat
atatgttaatgtgttttccaaagcctggcagacctgggttttgagctttgtatcacatgttttatgtttg
gaatgaaaatgagaccatgtctgtgaaggcactttgatatgcgtaatgcactctgccagtggtttgtcaaac

[illegible]

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TTTGACAAGTCTGTGCAATGGATCAAAATCAGAAGAGATGATTCAACTTGGTGACCAAGAAGTTTCTGAGCT
TTGTGGCCTACCAAGGGAGAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCCTGAAGCC
AATCCTGgtgagtagacttgctcactggagaaaacttcaagcactaatgctttcgggaatgtgaggcttttccct
tggaacagcatgactttgtttttagtaaaaagtacggctggctgggagtttgtgatataatttagttcagtggt
attctaagtgttcttagtggttctttcagacttttgggccatctcccaaagggtgaatgggaagaataagctg
gggtgtggctgagtttaagccaaaagttttttgtgcttggttcaatcagagaagacctgctttttcatgtttt
tactattataataactaagcaagagctcatttgaaaacagagttcttcatatttaaaaaaaaaaagtcttgaa
accattgatgggaagatggatatctattttagtttaaaaacccatcataaagatgacattgtgggctgtcac
agttggaagggccctggaattagatgagaccacactatttagcttacttagtaataacattgcaaagaaaaat
tccgacgaagttttttcagcctaggaatcaatagttcagagaagcactctatgagaataccattcattctt
aaccaaaaaataactgggtgagcctgagcagtttgggtcatcagagtgttttatatagttccagaacaaatatgt
ctctaggtgttctgagagctctgggtgaaattcctctcgtacccccaaacatcatcatttaatatccaggatt
ctggtttttctactcaccagatagattctcttaaaaccagggaagattcctggaggaaggatgtatctggaa
agagatgttcttattataataaaatgaaattgtaatactcttggattttgtgcagcagcaattctttatag
agagttgggtctccagagaattaagaatactcagtttctggaccctgttcccagatcataccctagaatgt
gaccttagaaacacacttcaggattcatacctttgattgaccatcaaaaagtttttgtatcggccagggtgtg
gtggctcacgcctgtaatcccaccactttcggatgccacggcgggcagatcacgtgaggtcagcagtttgag
accagcctggccaacatgggtgaaacctgtctctacaaaaatacaaaaaatagcggggcgtgatgggtgggca
cctgtaattccagctactcgggagggtgagggcaggaggatcgcttgagcctaggagggtggagggtgcagtgga
gctgagatctgtctcactctgttgcccagggtggagtgagtgaggcgatctzcgactcactgcaacctccg
tctctcatgcttaagtgattctcatgcctcagcctccgagtacctgggactacaggcacctgbcaccacgcc
cagctattttttgtattttztagtagacatagggtttcactatgttgcccagctgggtctcgaactcctgagc
tcaagtgatctgcccacctcgggctcccaaagtgttgggattaaaggcatgagtcaccgtgacctgggtccca
tgttataattttaagtaaggatatttctctacagggatctttgcaaccctaagtaazctggcctaaaaag
ttagagaagctgactgtgtgcagacatttgagcctgttgggtctttttgtgctgtgaattagaggggtgaa
agggtattatgaatgggtacaaaactttgttacaaaaccttttcttggactgttttgggctgcttccactgca
tgacaaatgctcaccctttcagctggaatgattgaaattttggaaaagatgggtgtttttagaagacatttgt
aatttgttccgggtgctgtgcccattcattccatttcaactctgtttactcattaaacacctatttgtgtacac
aaccgggtaaaatccctccactcacacaatgctgaattatactcatagtagaatgactgttttagccctcat
catctgataattaacagctcagggtttcaacctgacagtatctctctgggaggattagcagcgtgacagagt
cagggaaatgcaccttcagaaccgtcagctacactgtgtcccatcctgctgtgttgtgtgtgtgtgtgtgtgt
atgcgttgggtttatgaccagggtattgattaagggtgggtactaccagggtgctttctgcatatctcgggttgt
ggagcactcagggtctgcttctgcccctctgctgttaccagagacctctcttcaaaatggggctcttgagt
tagagtagaatgagtgatcaggattgttttgtgtaaatgatttctgaggaaggctttaggatgaaatgac
ttccaaacattttgaaatgtgactcttacttattgaaatgaagcagggtccttaattggaatgctgggactgat
acttgatttgcattaaagcagcctttttctattgtctgcttgggttgaaatttcaacatttgtgatggtagatg
gatgtgacatgtgatgacattgcacatgggcagtttaactgtgccaagaagtgcagcagtagcagcaaccgga
gatgcaaagcccaacatgatggggagagaaaactcttctttcaatatgtgcttctgtaccaaagtggaaattt
cacgagagacatttttgaacatttctccttttgtgtgtgcgtgagtggttccctgtttccagccaagggt
attgtgagttctcctgggctccttcagaatctgggtgctctggaagcagtggttttggcaacatggggaa
agtagtggcagtggtgggagggtcagctgggtctgggtttgaaatttgcaatttgaatattttaccagcattgat
gtcggataaattatttagtccctgtaagcctcagttttctcttctctacatacacataatatatttgactc
tttgttgtgattatttggttacacatatgaagagcctgggtgtggggcctggcacacaatagggtgctcaataaa
tagaagttgataatttaattgacatgagtagtagaattatgtccttgaaaacaattgcgtcaagatagaag
ttttcagccaggcacagtggtcacatctgttgtaatccagcatatttgtggggggcggaggcgatgaatca
cttgaggccaggagttcaagaccagcctggccaacgtgggtgaaatccctctctactaaaaatacacatatt
tgccaggcaggcgtgggtggcgacacctgtaatccagctactgaagaggctgaggcacaagaatcgcttga
accaggagggtggagggtgcagtgagctgagatcactccactgcattccagccagcgtgacagagtgagact
ctgtctcagaaaaagaaaaaaagatagaagttttcttctgtagatcagtggtagaactcataccaagcgaa
gtggctcctgggtgagtagttttagtgaaaaactgcattcttctcagatatttgtcaagacttttccccaaaga
ttcttattttatgtctcagtcctgaccttgtgtgaaaattaatactggatgtcagaacgctgttgtgttttta
aagttccctggggttaagagcagtttccattagggtgttctctgcttttacttaaaaaatcttactcatgcat
tgagcaatatattttagttcttattatgtgtcaggatatttctaggagctggactcaactcaaaagatatc
cttttgatgagaacaaagggtgggtggatatataatgaaatattatctgtgggataaatgcacttagtcatgagg
agactgttatggagtgcgtcattgtatttgactgttgagtttaacaacttctaggaggagctcaggggca
cctggcagggtctcttttgtcttctgctcagcaagggtgtattttgtgtgagagtgatgggtgggca
cttttcttaactttctcttgggtccttccataagcagcatgtacctttccagagcaggagaggggcacct

[illegible]

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

agtttcaacatgaaaaaagggtacagagaataacataaagaacactcctggctgggtgtgggtggctcacgcct
gcaatcccagcactttgggagtgctgaggcagccagatcactggaggtcaggagtggagaccagcttggcca
acatgggtgaaacactgcctctactgaaaatacaagaattagccaggcatgggtggcgtgcacttgtaatccca
gctacgtgagagactgaggcaggagaattgcttgaaccaggagggggagggtgcagtgagctgagatcaca
ccactgcactctagcctgggtgacagagtgagactccgtcttcaacaaacaaacaaacaaaaagaacactc
ctgtaccatcatccatcattttgccgtgctgactccagggttctatttaagaaataaaacattacagggtacag
ctgatgccacctctgtttccctagctcattcttcagagataactcttgtcttgcagttggatgttttaatacc
tctatatcaztgtatacttacattctatgtataacaatatttgggtactggcctaaatgtgttcacattgtat
aagtgtgcataattggcctgccacttcatttgggaattatgttcttgagatttatcaatgttgatacatgtgga
atctgggttaatttttggccatagttatctatttttataactaaacttttaaaaatccatgcttctagtctttgg
cttattttttcagggttatgggtatgttttggatgcacagaaaagttaaatttaagtcatgagcaaaatatctgg
ataatccaagcttttaaacttgatgtagaatttgaatcatgtgtgttttgttaacctgtgatgtcaatccat
gcctgatgtgtgtaactccaaccaatattcctttgaaaatggaaatttgttttatattgactacagattgcca
tattattagtaaatgctgagcacttaatctcgaataaagaactagtttaaaaatgatttcaacaatggcatt
gactgttctaccttattactcatgggtgggttcagccaatgtttctgttggagaccaaaacaaaaacagtc
aaatataacaagcagtcacaaaccaacatacagactactgataagaaggatcatatcataagatatggcattga
atttgtgtctgctaattgtaaaaatctgatgccacagcaacttaataaggacctatgtttacatccatgct
caattacattcctgggttaaacagtcagcttttaggcctgctgtgtgcctggagttttgcagagtgtggg
gcttttaagagaaggagaataagcttgcctcagagttaagaaatttaaaactaaaagtcctaaagatgttggg
aaaactattgccccttgaagatgtaaattcatttaagtgggagaagaccttttatacaaaacacagacccattc
actgatttgtacccttcaggagacagatgaccggtaattgggtgacaatgggtgaatgttgggtttgggggttt
tagaaacatctgcacttgggtgactactgtatctaattgggtgtgacaaacctggcaccocatgtgtttggcac
catcttgggtcctactcagggccagggtgaaccagagtggcctcttactgcttcagagtctgaagcagattgt
agtatgcccagcagacacagaatataccaccaagcacgttggcgcaaagcataactgggaaggagggtttgt
gaacatgggtgctgggttctcaaacctcagtgctcaaaagagtctcctgaggattcctggatcacactcttaaa
cttctcattcagtaggtgttagctgggtgagaatctgcattttttttttttttctgagaccaggtctcact
ctgatgccagggctggagtgtattggcgccatcttgggtcactgcaacctctgcctcccagggtcaagcaat
tttctgctcagcctctctagtagctgggttacaagcacatgccacctgctggctaatttttgtactt
ttagtagagaggggggtttcgccgtgttggccaggctcgtcttgaactcctgacctcagggtgatccaccact
ttggcttcccaaagtgttgagattacaggtgtgagccaccatgcccagcctgaatctgcataatttaacaagc
accacaggtgattctgatacagtagctccccaaacctcacagtggttagtgaatcccagtcatttacaattct
gccatgatttttggtcataattcaagtgcagctggtagcatttttagttaatatatttttaaatgaagtcact
tcttttggataattaaattttaattacaagggaagctataccactgctgttaaaaacatcacctgctttaaga
gaagggtacataatgaatatacattaaagataaagatgtatatgtgtgtgtgtgtgcacataatgaatatacaca
tacctaccatagggattgagttttccttcagggttttcaactgaaatgtcaactttgaggccaggttaatatg
tgtaagatatatgtgtgtatgtatgtctatacatatagacataacgtaaaaacatacatggatgcataatag
tatactatacacaacctattatgcataatcatgtatatttcatccacttagtattatcttztatttttgcog
tttggcaaatgctcagtaaaagaaaagggttagaaggggagaaaaggcattttatcccaagccttcagggaatc
aggatgaggatgtcttcaccttgtgggtggggagtazattatacaattagagacagcacattggagztgtggc
tgatatgctgtgtgatgatagctctagctctctgcctagcagaggaaggacatttcaatagaagaaaaagtt
taagaccttggcgagaaacagagaaaaggatgtttgtctttttaagaagttgaaaaccctgtttgcagacaaa
agccctccagtttttggcagtaaaacttcatgcaagggaagaaaaaggcaggggatgacattgttgacaattg
tgagggaattaccatgtgccaggcactgtgcgaggggtttgtacatatcctctagttttagtgcttataaaa
actctgtgatatgtgcacagcattttaaactttgctgcatagtcgagaaaatggaaggatgggggaatttgag
tcatttggccagggttctatagctaccccagggttcccatgactggagaattgggggcacagggtggcgggggga
gagtgagtgaacaagaatcctaacaatcttatttccattgagtccttataaaaagaagtggattaaactaccag
tttttaagtttttcttaaattttaggttatgtggatctggcgtttcttgttttgcctgggtttgttttgttt
ttgctatgtctgttgaacatctgtcatctttaggcctaacggtaaacacaaaaaacattacctcctata
gctttcaattaaagatctctcagtttgtgttggtaatagttttccaggcaagttctccctagggttcggcttct
agtggtgttaaccttttagttataaagtgaacccaagagagaaaagtagaaacaaaacacctcacctgttttgc
ctcatgaattactctctatggaagggaacaatcatgaacacctctgcgtatcacagaggcctatctgagctgc
acgtttaaggagagaccgcgtagggtccctttgaggactgtgaatgtgggagtcctgggaactctgggtgaagaac
ccgttccagaagagatgaatgagctggacaagttctttcatagaaccttttaggcagggttttcttagaaatgc
acattgaggattatgcttggatattgtgatgatcagaatgatactcaatcccttctgcatttgggaattctct
ttgaaagaaaacatcccaggcagctatttctcagagatagtgagtcacagccacttctagacattttcttgc
gtagctacattataatttccagcagctctctgatatgacaaatgtcaaaatagcccaaccttctctaaact
tcagagatgtctgatgatgatattgaataaaacaatgctcatagaacatcaagaaagggtggattttccctgg

[illegible]

gattttttttataaagacagcaggtttttcagaccctttggagactccaattcggtagaacacagagcttcatctctct
ctgtcgaaagctgtgacaggaggttgcaaatgcctctcctttttgctgagtttgacagctgctgtttttccggca
gcacatctgtgcaggcctctgcctcgccctctggzatctgctgattgagcagcggattgatctgtccttc
tctttcgtgttgacccatgtgaggaaccaactggcaagggazacaagaaatggaaataggcctcctttgcat
catgacctgtacatcctgcaattggaaaagattgtacttttagttgggttaaccagcagcattatttttctaa
actaagcagtaagaaggaattagggttttatgtgggatcaacagactgggtctcaaaagaggaaggtgataga
acacagtggggagggggaggtgcactagaaacagagggcctatgctttcattctggccttgctacttaatag
ctgtgtgaccaaatcttagagacttaacctctctgaacttccattttctcatgtataaaatgggaaatatta
aaggatactcactgggctgggtggcttgtgcctgtaatcccagcacttggggaggttgaggtgggaggatcac
ttgagcccaggtgttcaagaccagcccaggcaacatggcaagactctgtctctatgaaaaattaaaaatt.
gccaggtgtgggtgggtgtgcacctgtagtcttagctacttggtaggctgagatgggaggatcacttgggcttg
ggaggtcaaggtcggtgagctgtgattccatcactgcactccagcccggggcgagagcagagacactgaa
tccaaacgacaacaacaaaggcacaataaaagtgcctccttttaggttggtgtaaggtgaagcagaagc
tatacactattcaacatgtaactatataaaggaagtattgctgtgttactgtagttaataccattaaagt
agatgtttcgtatagtggaaagcacatggactctgaattcagactgggtctgactttgagtctcagctccaca
tctagtaatactatgaccaagccctgggttaaaatcatgttttttttttcttcagcctcagctctctcacatat
aaaaatagggacactgtcattttacctcagttttctgtgaggataaaaacaacgacagtgatatgcaagtattt
tgtaaattttgtagtgtcctcaagatttagttgggtgtttactacttgtactttctcactggaatggcagAT
GCTGTTGGACAGCAGGGACAATGACCACTTTTGGGAACAGCAGTTGGATGGCTTAGATTGGACAGCCCAAGA
CATCGTGGCGTTTTTGGCCAAGCACCCAGAGGATGTCCAGTCCAGTAATGGTTCTGTGTACACCTGGAGAGA
AGCTTTCAACGAGACTAACCAGGCAATCCGGACCATATCTCGCTTCATGGAGgtgaatctggttgctgggatc
atthagaaaagacttaacggcttctttctctgagacgttacaataaggttcaggcaggagggaagtttagaa
ataatgtatagtctcatttacaaaactatccctcaagcctaacacaggatttgataacaaaaggcacttaat
aatgttagttgagtgggtgaatgagtaataaactctagcttttagtaaatctactctagcttattctatat
aggctcaagagaatattttctacctttttctctaggttttctctatctcagtgactaatggtagcaaagcat
tcccttaaaaaaggcattatttgtgaaacttatctaaaatcgaattcgggtccaattaaatttttgaaatttt
atattaaaaattatattagtagggatgggtgaagaggtgttttgggtctgggtgggttagttgctatgact
cagaattgtctaagaaaacgaaaagtaagataagatcattgttttaacctttttctccacaaaatcaata
ataaacatattccctaaattactcttagaatttctctaaattgcagtgaaaaaccaaaatccttcattcttg
gttgaaggttggaaaaactacgttagagaggattagagagagaggatgagcaatcgtgtagtgcagccctggc
tctagtgtaggatttgtctcagccactgcttgtgtgctctggctgccaacgttctcatgaaggctgttcttc
tatcagTGTGTCAACCTGAACAAGCTAGAACCCATAGCAACAGAAGTCTGGCTCATCAACAAGTCCATGGAG
CTGCTGGATGAGAGGAAGTCTGGGCTGGTATTGTGTTCACTGGAATTACTCCAGGCAGCATTGAGCTGCCC
CATCATGTCAAGTACAAGATCCGAATGGACATTGACAATGTGGAGAGGACAAATAAAATCAAGGATGGgtaa
gtggaatcccatcacaccagcctgggtcttggggaggtccagagcacctattatattaggacaagaggtactt
tattttaactaaaaatttggtagaaatttcaacaacaacaaaaaaactcaacttgggtgtcatgattttgggtg
aaattgggtacatgacttgctggaaggtttttcataggtcataaaaataacagtatcttttgatttagcatttc
tactcaaggggaattaattccaggaattttgggtggcaggcacctgtaatcccagctactcgggagggtcaggc
aggagaattgcttgaaccaggaggcagaggttgcaagtgcataagatcgcatattgcactccgcctggg
caataagagtgaaactccatctcaaaaaaaagatacaaaaaatgaaaaaggggcttggtgaagggtag
taggggttttgggcaatt
tgggtttgggttttagGTACTGGGACCCTGGTCTCGAGCTGACCCCTTTGAGGACATGCGGTACGCTCTGGGGG
GGCTTCGCCCTACTTGCAGGATGTGGGTGGAGCAGGCAATCATCAGGGTGCTGACGGGCACCGAGAAGAAAAT
GGTGTCTATATGCAACAGATGCCCTATCCCTGTTACGTTGATGACATgtaagttacctgcaagccactgttt
ttaaccagtttatactgtgccagatgggggtgtatatatgtgtgtgcatgtgcatgcatgtgtgaatgatct
ggaaataagatgccagatgtaagttgtcaacagttgcagccacatgacagacatagatatatgtgcacacac
tagtaaacctctttccttctcatccatggttggcacttttatctttttatttttttttttttttttttttttt
gagtgctcgtctgacgcccaggctggagtgcagtggctcgatctcggtcactgcaacctttgcctcccggg
ttcaagctatttctcctgcctcagcctccacagtagctgggactacagggtcatgctgccacgcccgggtgac
tttttgatttttagtagagacgaggtttcccatgttaccagggtagacttcaactcctgagctcaggcaa
tccacctccttggcctcccaaagtgtctgggattacaggtgtgagccactgcaccagcccaccactttaat
tttttacactctaccttttgggtcaaaatttgtcfaatctgcaagcttaaaatgtgtcatgacaaacacatg
caagcacatactcacacatagatgcagaaacagcgtctaaacttataaaagcacagtttatgtaaatgtgtg
cacttctttccctaggtggttaaaccacatttcaaaacaacccaaataaaactgaacaaagcttcttctct
tagacttttttagaaaactcttcagtgctgagtcaactaagctgccaagttctcattgtgggaactatgccttt
ggatgtaatgattttctctcaagacaatggggcggaaggttagttattgcagacatctgaaattatgtaatgttt
cttccagattctggaaattctcttattctctgtggttgggtgggtgggtgggtgggtgggtgggtgggtgggtg

[illegible]

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tactttcattgcaaattggaaacttcgtttaaaaaatgctcataactaaaattggcatttctaacccatagggc
ccacttgtagttattttaccgaagcaaaaggacagctttgctttgtgtgggtctggttaggggttcattagaaag
gaatgggggcggtgggagggttgggtgttctgttctctctgcagactgaatggagcatctagagttaagggtta
gggtcaaccctgacttctgtacttctaaattttgtcctcag**GTCAATCCTGACCGGGTTGTTCCCCCGACC**
TCGGGCACCGCCTACATCCTGGGAAAAGACATTGCTCTGAGATGAGCACCATCCGGCAGAACCTGGGGGTC
TGTCCCCAGCATAACGTGCTGTTTGACATgtgagtaccagcagcagcgttaagaataggccttttctggatgt
gtgtgtgtcatgccatcatgggaggagtgggacttaagcattttactttgctgtgtttttgttttttctttt
tttcttttttatttttttgagatggagtctcgtctctgtagccaggctggactgtagtggcgcatctcggct
cactgcaaccttggcctcccaggttcaagcgattctcctgcctcagcctcccaggtagctgggactctagggc
acacaccacctgcccagctaatttttgtgttttttagtagagacgggtttcaccatgttggccaggatgggt
ctcaatgtcttgacctgctgatccgcccacctcgggtctcccaaagtgtgggaacacaggcatgagccactg
tgtctggccacattttacttttctttgaatatggcaggctcacctcgtgaacaccttgagacctagttgttc
tttgatttttaggagaagtgggagggtgaatgggttgagctgtagaggtgacatcagcccagccagtggtgggg
gcttgggaaacattgcttcccatattgtcatgctggagggtccctttagcccatcctctcccccgccaccc
tcttatttgaggcctggagcagacttcccagacctggtagtgcttcagggtccctgggtatgatggacctatat
ttgtgtcttaagacattttgctcccactcaggttgtcccatcagccataaggccccagggtgagccgtgtgat
ggagcagagagagacctgagctctgcaatcttgggcaaggcttttcccttatgtttcttcttatctaaagt
aacagctggggctcatgtgtctcctcctcatctaaagtgaacacatggggctcatgtgcagggtcctccccg
cttcagagcctgagggtccctgagggtcaggaaggctgtctccagggtgagtgcgagctgacttcttggtgg
acgtgtgtgtggggacagccattaaagaccacatcttggggccctgaaattgaaagttgtaactgctgggtg
catggtggccaggcctgtctggaaacagggttggaaagcagctgtctacaccttcaactttgatttctgagcagct
catgtggttgtcactgttgttctaccttgaatcttgaagattatttttcagaaattgataaagttattttta
aaaagcacggggagagaaaaatatgccattctcatctgttctgggcccagggtgacactgtattctggggat
ccagtagggcccagagctgacctgcctccctgtccccag**GCTGACTGTGAAGAACACATCTGGTTCTATGC**
CCGTTGAAAGGGCTCTCTGAGAAGCACGTGAAGCGGAGATGGAGCAGATGGCCCTGGATGTTGGTTTGGC
ATCAAGCAAGCTGAAAAGCAAAACAAGCCAGCTGTGAGgtgcccagagctaccttccctatccctctcc
cctcctcctcgggtacacacatgcccaggaaaatcagcactgcccagggtcccagggtgggtgcccgttgg
taacagaaacttgtccctggctgtgcccctagggtcctctgccttcaactcactgtctggggctgggtcctggag
tttgtcttgtctgtttttttgtag**GTGGAATGCAGAGAAAGCTATCTGTGGCCTTGGCCTTTGTGGGGGA**
TCTAAGGTTGTCAATTCTGGATGAACCCACAGCTGGTGTGGACCCTTACTCCCGCAGGGGAATATGGGAGCTG
CTGCTGAAATACCGACAAGgtgcctgatgtgtattttattctgagtaaatggactgagagagagcggggggct
tttgagaagtgtggctgtatctcatggctaggcttctgtgaagccatgggatactcttctgttakcacagaa
gagataaagggcattgagactgagattcctgagaggagatgtgtgtctttattcatctttttgtccccaac
atgggtgcactaaatttatgggttagttgaaaggggtggatgcttaaatgaatggaagcggagaggggcaggaag
acgattgggctctctgggttagagatctgatgtggtacagtatgaggagcacaggcaggcttggagccaactc
tggcztgggccctgagacattgggaaagtacacacttgccctcaccttctttgccgataataatagtggtgcz
tacctcatagaggattaaatataatgagaatgcacacaaaccacctagcaccaatgctggcatatagcaagt
tcccaataaaaatgcztactgttcttacctctgtgaggatgtggtacctatataacaaagctttgccattc
taggggtcatagccatacagggtgaaaggtggcttccagggtctcttccagtgcttaccctgctaatactc
tctagtccctgtcactgtgacaaatcagaactgagaggcctcacctgtccacacatccttgtgttgtgctctg
gcag**CCCGCACCATTATTCTCTCTACACACCACATGGATGAAGCGGACGTCCTGGGGGACAGGATTGCCATC**
ATCTCCCATGGGAAGCTGTGCTGTGTGGGCTCCTCCCTGTTTTCTGAAGAACCAGCTGGGAACAGGCTACTAC
CTGACCTTGGTCAAGAAAGATGTGGAATCCTCCCTCAGTTTCTGCAGAAACAGTAGTAGCACTGTGTCATAC
CTGAAAAAGgtgagctgcagctcttgggtgtctgggctgggtgttgggtctgggcagccaggacttgcgtggctgtg
aatgatttctccatctccaccccttttgccatgttgaaaccaccatctccctgctctgttggccctttgaaa
tcatatcatacttaaggcatggaaagctaaggggcccctctgctccattgtgctagtctgttgaatcccgt
tttcccttttccctatgaggcacagagagtgtatggagaaggctccttagaggacattattatgtcaaagaaaaga
gacttgtcaagaggtaagagccttggctazcaaatagacctgggtzgttccctgctcattacttttcaatctcat
tgaccttaacttttaactataaaacagccaatatttattaggcactgatttcatgccagagacactctggg
caztgaaagaaagtaatgataatagtttaattttatatagcgttgttaaccatttacaaccttttttttttt
taacctctatcatctcaattaaagtgcagagagaccttgggaagaaggtaactatatttattatcccagatg
agggaagtgaggcttgttagggaattggtagctgattcaagggtcacccagcaggtaataaacagtgggtgggac
cagacccaattaccagggtatgttttccctctgtaccgcagtagcatgctgagatttatttgtgtgttgaagcc
agtggtagcctaattgtatttacatcccaacctgaaactcctatccacttatttaccttttaatgagcctctta
actcaagtgcagctctgaggaccagcagcatcaggatcacttgggaacttgttagaaattcagcaacctgggc
ccagctcagacctaccgaatcagaatctgtgcattttaacaagggtcttgagtgggttgaacacacattaaag
catgagaagcattgaactagacatgttagccagggtaaaggccttgccctgagatgggttggcaaggccctcattg

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cagcattcattggcagggccacagttcttttggcagctctgcttctgacctttcaccctcaggaagcgaggc
tggtcacacggccacacacatgccagacaggggtcctctgaagccacggctgccagtgcagtggtcccaggga
agcttttttcttttagttctcacacaacagagcttcttgggaagccctccccggcgaaggtgctggtggtctg
ccttgctccgtccctgaccgttctcacctccttcttggccatcag**GAGGACAGTGTCTCTCAGAGCAGTTC**
TGATGCTGGCCTGGGCAGCGACCATGAGAGTGACACGCTGACCATCGgtaaggactctgggggtttcttattc
aggtggtgctgagcttccccagctgggcagagtgagggcagaggagagaggtgagagggtggtggcg
tgactcaaggtttgctgctgggctggggctgggtggctgctgggkgtgaggagcagcttgggtggcgggttggcc
taatgcttgctggggtgctggggctcgggttgggagctagcagggcagtgctccagagagctgagatgatt
gggggttggggaatcccttaggggagtggaactgaataaccagggtgaggagctgagggccaagccaggag
ggtgggatttgagcttagtacataagaagagtgaagagccaggagatgaggaacagcctccagattttct
tgggtagcgtgtgtaggagggcagtgctcaccagtagcatatgtggaacagaagcttggaccttgctatctc
tgacctagtcctaagtggtggcttttccagggaaggtcttgcttccatggactgttagattaaccctttatt
taggtaaatgaggggaacctactttataagcataggaaaggtgaagaatctttaagattcctttactcaag
ttttcttttgaagaatccagagcttaggcaatagacaccagactttgagcctcagttatccattcacccat
ccaccacccacccacccatccttccatcctccatcctccattcacccatccaccatccagctgtccac
ccattctacactgagtacctataatgtgctggcttgggtgatacaaaggtgaataagacatagtcctttcc
tttggccccaacccctcagaccagagatgaacatgtggaatgacctaaacacctggaacaggtgtggtgatg
agcggcagggcctctgatgagaggggtgggggatggccagccctcactccgaagccctctgagttgattgagc
catctttgcattctgggtccctgcag**ATGTCTCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCC**
CGGCTGGTGAAGACATAGGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGGGAGCCTTT
GTGGAACCTCTTTCATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTAGTTATGGCATCTCAGAGACG
ACCCTGGAAGAAgtaagtttaagtggtgactgtcggaatatatagcaaggccaaatgtcctaaggccagacc
agtagcctgcattgggagcaggattatcatggagtttagtcattgagtttttaggtcatcgacatctgattaa
tggtggccccagtgagccatttaagatggttagtgaggatagcaggaaagaagtggttttctctgtaccaca
gtacatgctgagatttgtgtgttgaaaccagtggtacctaacacatttacatcccaaccttaactcctat
gcacttatttaccctttaatgagcctctttacttaagtacagtgkgaggaacagcggcatcaggatcacttg
ggaacttggttagaaattcagcaacttgggcccagctcagacctactgaatcagaatcaggacaattctctg
gtgtgactgtgtcacagccaggtatcaactggattctcatacataggaaatgacaaacggttatggatggat
agtctacttggtgccaggtgctgagatttgtttttgttttttgatttttttttaactactgtgacctcattt
aattctcaaaaaaagatgaaaaaatgaacactcaggaatgctgacatgagattcagaatcaggggttgggg
cttcaaggtccatcctctctttatccatgtaatgcctcccttagagatacaacatcacagacctgaaggc
tgaaggggatataaaaagctgtctggccaagtggtctccaagcttgacagtgcagcagaatcacctggggata
ttatataaaaataaacataactaaggttgggttcagggcctgtgaatcagaatttctggaggtgaggccttga
agtctgtatttctattgcatactttggacacagtggtctatagactagagtttggaaatgattgcgctcatt
cagattctctctctgatgtttgaattgctgccatcatatttctagtgtctatttccctcctgctcattctgtc
ttggataacttatcatagtactagcctactcaaagatttagagccacagtcctgaaagaagccacttgactc
attccctgtaggttcagaataaatttctctgctgcagtgctgtcatagcttttttaaattttttttatt
tttgatgagactggagttttgctcttattgcccagctggagtgagtggtgctgatttgggtcactgcaac
ctccacctcccaggttcaagcgattctcctgctcagcctcccaagtagctgagattacaagcatgtgctac
cacgcccagctaattttgtatttttagtagagatgggttttatccatgttgggtcaggctggtctcgagctcc
agacctcaggtgatctgcccgcctcggcctcccaaagtgtgggattataggcctgagccacagcgctcagc
cataactttaatttgaaaatgattgtctagcttgatagctctcaccactgaggaaatgttctctggcaaaa
cggcttctctcccaggttaactctgagaaagtggttattaagaaatgtggcttctctctgtcttacggg
gctaacatgccactcagtaataataaatcgtggcagtggtgactactctcgtaatgttgggtgcttataatg
ttctcatctctctcattttccag**ATATTCCTCAAGGTGGCCGAAGAGAGTGGGGTGGATGCTGAGACCTCAG**
gtaactgccttgaggggagaatggcacacttaagatagtgcttctgctggcttctcagtgccagagtattg
ttcctttcccttgaattgttctattgcattctcattttagagtgtaggtttgttgcagatggggaaggtt
tggtttgttgtaataaaaataaagtatgggattcttctctgtgcttccag**ATGGTACCTTGCCAGCAAGAC**
GAAACAGGCGGGCCTTCGGGGACAAGCAGAGCTGTCTTCGCCCCTTCACTGAAGATGATGCTGCTGATCCAA
ATGATTCTGACATAGACCCAGgtctgttagggcaagatcaaacagtgctcctactgttgaatgtgaaattct
ctctcatgctctcacctgttttcttggatggcctttagccaaggtgatagatccctacagagtcctaaagag
aagtgaggaaatggtaaaagccacttgttcttgcagcatcgtgcatgtgatcaaacctgaaagagcctatc
catatcacttcccttaagacataaagatgggtgcctcaatcctctgaacctatgtatttattatctttctg
cggggctcctagtttcttgtatacattaggtgtttaattgttgaacaaatattcattcgagtagatgagtgat
tttgaaagagtcagaaaggggaatttgtgttagagtttaattgtaccctaagacttagatatttgaggctgg
gcatggtggctcatgccagtaatcccagcgtttgagaggctgaggtgggttagatcacctgaggtcaggagt
ttgagaccagtcctgaccaacaaggtgaaaccccgctctctactaaatacaaaaaattagccagtggtggtggc

acatgcctgtcatcccagctacttgggaggtgagggcaggagaatcgcttgaacccaggaggcagaggttgc
agtcagccacggttgcgccattgcactccagactgggcaacaagagtgaactccatctcaaaaaagaaaa
aaaaagaattagatattttggatgagtggtgtcttgggtgtttaaactgagatggagaggagagctaagacat
caaacaaatattgttaagatgtaaaagcacatcagtttaggtatcattagtttaggacaaggatttctagaaa
atTTTTtaggaacagaaaactttccagtttctctcaccctgctcaaagagtgtatggctcttacattatata
aactgcctgacttcatacagtatcagtttagatcatttgaatgtgtccaggttttaccaaaatataata
gggtgagaagctgagatgctaattgccattgtgtattctcaaatatgtcaagctacgtacatggcctgttct
atagagtgtctataaagaattgatgacttgattcatccgaatggctggctgttaacacctggttacgcatga
acacctcttttcagttgtctcaagacacctttctttctgtacttatcagacaaggactgaaaggcagagac
tgcactgttagacattttgagtcagcttttcttggacatagcttgtcatgaaagcccttacttctga
gaaacttctagcttcagacacatgccttcaagatagttgttgaagacaccagaagaaggagcatggcaatgc
cgaaaacacctaagataataggtgaccttcagtggttggcttcttgcag**AATCCAGAGAGACAGACTTGCTCA**
GTGGGATGGATGGCAAAGGGTCTTACCAGGTGAAAGGCTGGAACTTACACAGCAACAGTTTGTGGCCCTTT
TGTGGAAGAGACTGCTAATTGCCAGACGGAGTCGGAAGGATTTTTTGTCTCAGgtgagacgtgctgtttcg
ccagagactctggcttcatgggtgggtgcaggtctgtgaccagtgaaggcaggatagcatcctgggtcaag
atatggatgccggagccagatttatctgtatttcaatccagttctattccttggcagttgtgtatccgctg
gcaagttacttctctatgcctcaatctcctcatctgtaaaatggggataataatattacctgcaatacaggg
ttgttacgaaaataaaaaatgaataggtgcttagaatggggcctgacattagtaagtgttagttttgtgtgt
gtatatgttatttttattttggaggagaacataaaaaggacaaagtgtagaaaaactgggttgggtgtattca
gctgtcataacatgagagttgttatgccagatgcacttgacatgtgaatttattagaaacatgattttct
ctgagttgatgtttaactcaaactgatagaaaagataggtcagaatatagttggccaacagagaagacttgt
tagactattgtctgcatgtcagtggttgcagtgtaacttgcttagttagaaaggttaaatttttctactcta
taaaatcaagaaatatagagaaaaggtctgcagagagcttttcatttgatgatgtggatattgttaagagcg
ggagtttggagcatacagagctcaagttgaatcctgactttgctacttattggctatatgaccttgggcaag
ctgcttagtctctctgatcctcagttacctttgtttgttgatgatgaccattgataacacaaccataaata
tgacaacatagagatagttctcattatagtagttgtttatacagaattattcactcaatgttaattttctgca
ttgaaatcccagaacattagaattgggggcatattttgaatctttaagggttataaggaatacatttctcagc
aataaatggaaggagtttgggttaacttataaagtatacccaagtcatttttttttcagagaagatatgggt
agaaagtcttaggaggttgaagaaggaattggatatatttcttctgagactatcatgggagataatgact
atgggtgtccatgattggagccgttgcgtgtagagttgggttttattatagtgtaggatttgaatgggcatgt
gttctcagacctcagattaaaawgagaaaaactgaggccagtgaggagcgtgacttcacatgggtacacttgt
gctagagacagaaccaggattcaggacttctggctcctgggtcctgggttcatggcccaatgtagtctttctc
agtcttcaggaggagggaagggcaggaccagtggtctgagtcacctgaatgtgagcactatttacttctgtg
aacttcttggcttagtgctctgccaggtggccataacctctggccttgtgttgccagagaaaagggttagt
tttcaggtctccattgcttcccagctgccaagaatgccttgggtgcagcacagtcattaggccctgcattcctca
ttgccgtgctgggttgggtcggggaggtgggctggactcgtagggatttggcccttggccttgtttctaacact
tgccgtttcctgctgtccccctgccccctccactgcctgggtaaag**ATTGTCCTTGCCAGCTGTGTTTGTCTG**
CATTGCCCTTGTGTTTCAGCCTGATCGTGCCACCCTTTGGCAAGTACCCAGCCTGGAACCTTCAGCCCTGGAT
GTACAACGAACAGTACACATTTGTCAGgtatgtttgtcttctacatcccaggagggggaagattcgagcag
accaaagatgtttacgagggccaagggaatggacttcagaattacacgggtggaatgaattttactgctgcgg
ctcaggtccctgtataagctaatactgcatgcatagaacagcagcgaactaacctgaataataggccagtc
ttctgttgagcctttcagcctctctcctctcactactgttgtcaggaacagccacatgtgttttaggtg
aaataatccaccttgcaaaaaatccatgattaagttataaaaatatttggatttgtggagctgtgttttaatt
ctgtaactgagtcacagggcacactgtcaaagcatagaacctccagagacttgttttctgcaaaagtataatt
catgtaattatttatctattctgttatatttgggatgttaggtagtgtttgttcttttagataaaaaatcccc
cactctgtaacaatacatataaatcaaagaaaaggacaaaggatttttctgggtcttgttagcaggagctttc
ttcagtcctgaaagatttgtagacctgtagatgggggaactgtgtcagtgatacaaaagggaagcattttaa
aaaaaaaaagtatatatatatatatatatatatgtaatgtgaattggcctcttttctctaagzcca
cattttzcttctacatagttcagggttactttatttttcttccggtgctgacctgtattgcccgtg
gttgtggaacatagcatgtgtttgtgacctgtgcctgttatttttgtgtcttctagttgtgcatgcaaaagag
taciaaagtttcttggcctttcttggaaaatcctgcttgtctgtgcaaaagggaataattgtgaaagcacttt
tgaaatacttaatgagttgattttcttcaaatataaaaaaatatataaatgtatatgtgtatgtacatgtgt
gtacacatacacacctttatacacacagcccatttaaaacaagctccactttggagtgctctacgtcacct
gatgccgaatacagggccagagctctgagatccttctgggtgggttctgtgttttgttcatttctgttttaag
agcctgtcacagagaaatgcttccataaatgttttaattataaaaaacatttttatctctcgatgattgggtt
taatgaattactaagctggctgcctctcatgtacccacag**CAATGATGCTCCTGAGGACACGGGAACCCCTGG**
AACTCTTAAACGCCCTCACCAAAGACCCTGGCTTCGGGACCCGCTGTATGGAAGGAAACCCAATCCCgtgag

cggggtgggtcacacctgtaatctcagtatcttgggatgccgaggtggactgatcacctaaggtcaggaggt
cgaaaccagcttggccaatgtggcgaaacctgtctctactaaaaattcaaaaattagccaggcatgggtggc
acatgcctgtaatcccagttacttgggaggtgaagcaggagaattgcttgaacctgggaggtggaggtcgc
agggagccaaaattgcccactgtactccagcctgagcaacacagcaaaactccatatcaaaaaataaaatg
aataaaataacagctaatactagtcatacagtataactccagtgaaacagaagatttattaggtcatagtgaatga
tgggtgcttccataaaaatctcttgactacaaagaatctcatttcaatgtttattggttagatgttcagaataa
attcttgggaaagaccttggcttgggtgaagtgaattaccagtgccgagggcaggggtgaaccaagtctcagt
gctgggtgactgagggcagtgcttgggacctgtagtccaggttccgggtcacactgtggacatgggtcactgtt
gtccttgatttgttttctgtttcaattcttctctataaagaccgtatgcttgggttttcatgatgagagaagA
GAAACAAAACACTGCAGATATCCTTCAGGACCTGACAGGAAGAAACATTTCGGATTATCTGGTGAAGACGT
ATGTGCAGATCATAGCCAAAAGgtgactttttactaaacttggccctgccktattattactaattagagga
attaaagacctacaaataacagactgaaacagtgggggaaatgccagattatggcctgattctgtctattgg
aagtttaggatattatcccaaactagaaaagatgacgagagggactgtgaacattcagttgtcagcttcaag
gctgaggcagcctgggtctagaatgaaaatagaaatggattcaacgtcaaatttggcacttagtagcaactt
gaccaggttaactgggttatccttttaaagccttagtttatctaaattgtgatattaatgttgcctctataagt
ttgtcatgaggactaaattaaatgggtgtacatagagtgccttgggtactctctgatgggggactccatgata
atttgtgggtctcatggaggagctctgggaaggtttaggagcctgccttggctctgcagccttgggagagcc
ttctagcttcccaggacatggcagcctagtgttgaatgcttggctcagcaaatgtttgttctcgtttccttc
ccatcaacttgggtcagttgggggtctttcagttaggagtatctcagtgactttaaatggcatgggcatgctgg
agtgatagtgacctagatttctaaagaaagaagcataatttctccatatgtcatccacaattgaaatattat
tgttaattgaaaaagcttctagggcaggcacgggtgggtcatgcctgtaatcccagcactttaggagggccaag
gcgggtggatcacttgaggtcaggaggtttagaccagcctggccaacatggggaaacctgtctctactaaa
aatacaaaaataagctgggctgggtgggtgcctgtaatcccagctacttgggaggtgagggcaggagaat
tgcttgaatctgggagggcggaggttgcagtgagctgagttcatgccattgcattccagcctgggcaacaaga
gcgaaccatctcccaaaagaaaaaaaagaaagaaagcttctagtttgggttacatcttgggtctataag
gtgggttgtaaaattgggttaaccaaggcctgggtctcatataagtaataagggtattctgatggagagaag
gctggaagaggcctgaacacaggtctctttctctagcacaccctacaaggccagctgattctagggttat
ttctgtccgttccctatatctcaggtggatatttactccttttgcacatttaggaataggctcagtgcttt
ctttgaactgattttttgtttctttgtctctgcagCTTAAAGAACAAGATCTGGGTGAATGAGTTTAGgttaa
gttgcctgtctttctggcacgttttagctcagggggaggatgggtgttgtagggtgtzcttggattgaagaaagcc
ttggggattgtttgtcactcacacacttgggggtgccatctcactgtgaggaggacagaagccctgtgaaca
tgtggagcacacaggggacagacagatttagattaggcctgctttatagagtttctgcctagagcatcatg
gctcagtgcccagcagccctccagaggcctctgaaatatttgatatactgatttcccttaggagagaatcaga
aatctcctgcaggtgtctagggatttcaagtaagtagtggttgtgaggggaatacctacttgtactttcccc
caaaccagattcccagggcttcttaaggactcaaggacaatttctaggcatttagcacgggactaaaaaggt
cttagaggaaataagaagcgccaaaacctctcttgcactgtatttcaaccatttgtccttctgggtttt
gaaggaacaggtgggactggggacagaagagttcttgaagccagtttgtccatcatggaaaatgagatagg
gatgtgggtacgtcagggggcccgaggctccttgttactgatttccgtcttttctctctgccttttcccc
agggccaggacccctggatctctgggcagagcagacgcagggccctataatagccctcatgctagaarggag
ccggagcctgtgtataaggccagcgcagcctactctggacagtgcaggggttccactctcccaactcccc
ctgcttgcctccagacccacattcacacmcgagccactgggttggaggagcatctgtgagatgaaacacat
tcttctcctaagtgtcagctatctaactgtgtgtgtaatcaggccaggtcctccctgctgggcagaaacca
tgggagtttaagagattgccaacatttatttagaggaagctgacgtgtaacttctzzgaggcaaaatttagccc
tctttgaacaggaatttgactcagtgaaacctgtacacactcgcactgagtcgtgctgctgatgatactgtg
caccccaactgtctgggttttaatgtcaggctgttcttttagGTATGGCGGCTTTTCCCTGGGTGTCAGTAAT
ACTCAAGCACTTCCTCCGAGTCAAGAAGTTAATGATGCCATCAAACAAATGAAGAAACACCTAAAGCTGGCC
AAGgtaaaaatatctatcgtaagatgtatcagaaaaatgggcatgtagctgctgggatataaggagttagtggc
agggttaaaccggtacacctggcagctcattgttctgaatatgttggcatacagagccgtcttgggcatttagc
gatttgagccagacaaaactgaattacttagttgtacgtttaaagtgtagggtcaaaaacaaatccagagggc
caggagctgtgggtcatgcctgtaatcctagcacttgggagggctgaagcgggtggatcacttgaggtcagg
agttcgagaccagcctggcctacatgacaaaaccccgatctactaaaaatacaaaaaaatttagctgggctt
gggtggcacacacctgtaatcccagctacttgggaggtgagggcaggagaattgcttgaacctgttaggaaga
gggtgttagtgagccaagatcgaccgttgcactccagcctgggcaacaagagcaaaactccatctcaaaaa
caaatataatccagagattttaaagctctcagaggttgggcgcgggtggcttacacctgttatccagcattt
tgggatgccgagggcgggcaagcacaaggtcaggagtttgagaccagcctggccaacatagtgaaccctgt
ctctgctaaaaacatagaaaaatttagccgggcagtggtggcgtgcgcctgtaatccagctactcgggaggt
gaggtgagagaatttcttgaacccgggagggcggaggttgagtgagccagattgcaccactgcactccagc

ctggcgacagacagcaagctaccctcatctcaaaaagagctctcagaacaaccagggtttacaiaatttgggtcaggtg
gtaaataaaactgggtttcaaacatactttgctgaaayaatcactgactaaataggaaatgaatctttttttt
tttttttttaagctggcaagctggctctgtaggacctgataagtactcatttctctgtgtctcaggtt
tcccattttttaggtgagaattaaggggctctgataaaacagaccctaggattgtggacagcagtgtagtcc
tagagtccacaagtctgcttttgagtgatgggcccattgtatctggcacatctgcaggcagagcgtgggtctg
gctcttcagatgatgccggtggagcactttgaggagtcctcaccaccacgtgataaccagacattaaaatct
tggggctttgcatcccaggattttctctgtgattccttctagacttgtggcatcatggcagcatcactgctgt
agatttctagtcacttgggttctcaggagccgtttatttaattggcttcacatttaatttcagtgaacaaggta
gtggcattgtcttccacagggccgtcctgttgtccacaggttcagattgactgttggcccttatctatgtg
aacagtcacaactgaggcaggtttctgttgtttacag**GACAGTTCTGCAGATCGATTTCTCAACAGCTTTGGG**
AAGATTTATGACAGGACTGGACACCARAAATAATGTCAAGgtaaccgctgtctttgttctagtagcttttt
gatgaacaataactccttatgtttctcgagtagctttcaactcatggttaaagttggcaggggcatcacaaca
gaaagagcaaaactattaactttaccagtagaggcagtcaggtgtagtgtagtgatcagagaatttgctttg
ccaccagacataccaggtaaccttgactaagttacttaacctatctaaacctcagttycctcatctgtgaaa
tggagacagtaatcatagctattttccaaactgttgtgagaattcaatgagttaaagggtataagggtcctcacc
acagcgctgcccacatagtcagtgatcactatgtcctgaacactgtaattacttcgccatattctctgatc
atagtgttttgcttggatgtgactagaatttctttctgaggtttatgggcatgggtgggtgggtatgcacc
tgcttgaggagcccggtttgggggcatcacttgtagctggatgttttctttcag**GTGTGGTTCAATAAC**
AAGGGCTGGCATGCAATCAGCTCTTTCTGAATGTTCATCAACAATGCCATTCTCCGGGCCAACCTGCAAAAG
GGAGAGAACCCTAGCCATTATGGAATTACTGCTTTCAATCATCCCCTGAATCTCACCAAGCAGCAGCTCTCA
GAGGTGGCTCTgtaagtgtggctgtgtctgtatagatggagtggggcaaggggagaggggttatgggagaagggg
agaaaaatgtgaatctcattgtaggggaacagctgcagagaccgttatattatgataaatctggattgatcc
aggctctgggcagaagtataagtttacgaattggctgggtgggtccttgaactgcagaagagaaaaatgac
actgatatgtaaaaatcgtaacatttagtgtaattcatataaaagttagttcaaaaattgtaattaaattata
atttaattataagtgtttaatcagtttgatttgtttaaaaaccactgttttaaatttgggtggaatatgtttt
tattagcttgtatctttaattcctaaattaaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
gtgtgtgzztgtgaagtttaaagccaggatgagctagtttaaagtagtgcagccttggagtcatacagatct
gggtttgaattcggtctctaaactttatagatgtatgatattaaatgaggcagttcatgtaaattgccaaagc
ccagcactcagcacagagttgatatttcacacacattagatacctttctgtatgtggagcatggcagttcc
gttttctgctttactcctacaggatactaataataggacactaggatctttataccaagacccccatgtaattgg
gcttatgagaccattcttcttataaaaaatctgacagaatttttgtatgtgttagatcaataggctgcatact
gttattttcaagttgatttacagccagaaatattaatttatttgagttagttacagagtaataatttctgtctc
catttagttttcaagccccactagtcccttgtgtgtgaaaatttacaacttactgtctttacaaggtcatga
acagtggaaccaaagtgaatgccattaaccactctgacttccttcattagttttattgtgacagtggaactcct
ttgacctcagtaataaccagtttggcatttacattgtcatatttttagacttaaaaatgatcatcttaacct
gaataaaatgtgtctggtgaacagatgtttttccttgggctgtgcctcagatatctctgtgtgtgtgacgt
gtgtgtttgtctgtgtgtccatgtcctcactgattgagccctaactgcatcaaagacccctcagattttcac
acgctttttctctccag**GATGACCACATCAGTGGATGTCCTTGTGTCCATCTGTGTCATCTTTGCAATGTCC**
TTCGTCCCAGCCAGCTTTGTCTGATTCTCTGATCCAGGAGCGGGTCAGCAAAAGCAAAACACCTGCAGTTTCATC
AGTGGAGTGAAGCCTGTCTACTGGCTCTCTAATTTTGTCTGGGATATGtaaggacacagggctgtgt
atctttctgatgtctgtcagggccatggattgatattgataagaagaagaagagctctggctatcatcaggaa
atgctccagctactctaagagattgtataaaaaagaatagccagaggcaggtgatcactttcatgacaccaa
cacagcatgtgggtaccagagttcatgtcacaccagaggggaaaattctgtacacaatgatgaaaattaatacc
actaccacttaagttcctatgtgacaactttcccaagaatcagagagatacaagtcaaaactccaagtcaat
gcctctaacttctctgatgggttttaacctccagagtcagaatgttctttgccttactaggaaagccatctg
tcattttzgaanaactctgtacattttatcagcagcttatccatccattgcaaataztgttttgtgccagcca
caatatattgtcttctattttggaccaatatgggggatttgaaggaattctgaagttctaattatatttcaact
ctactttacaatatctccctgaaatatatctccctgtaacttctattaattataagctacacagagcaaatc
taattcttctcccaccgaacaagtccttggaatttaaaaataactctcatactctcatttaacctgagtat
taccagataagatgatatatgagaatacaccttgtaacctccgaagcactgtacaaatgtgagcaatgatg
gtggagatgatgatgagatctttgctgtttataccaagcccccttagactgtgtcactcttctgatccggttg
tccttgtatggccatgctgtatattgtgaatgtcccggttttcaaaagcaaaagccaagaattaaaccttgtgtt
caggctgtggctctgaatggttatgggtccagagggagttgatcttttagctcacactcttacttactcagcac
aaagattttgcattttggaaggagcaccgtcttactggcaacttagtggttaaaccaaaacccattccattcaca
caaatgattgtgaaattcgggtctccttctattctatacaaaattcattgatttttttgaaactaaactttat
atttatccatatataaattacatgggtttttttttgttttatcttgattcagtaattactcctttcagtaaa
cacagactgaagtctgtgtgtgtgtgacttatgccaggcataggtgattcagagatgaaagggtcaagtcctgta

acccatctcttgtcttctcctgggtattatctgtccctccctgcttttagagctcctgaaatttgctagaagcat
 gtcttcatctaaagttgttgataaacacatcaagtaggattggactgaggcagagccctgtagctctgaagctg
 cagttcttctagcggctgacaagccccactatcacttccctgctgggtgctttgctctgccagctgtgaattc
 tcataattgtcctatcgtcaagtctttatttctgcattttactgcttgatacactgtcaggacagactttaa
 aattattctcagtgccgatgaaacaattctgacattcatgttatgagcagttacctcataaatagattacatg
 tgagattgaacttgggcagactataatatagcattaatgatgaaacagacacagtcattctcggaagaaga
 atagaggcttatttgctgcctgtgaaattaaaattactctgactgggaatccatcgttcagtaagtttactg
 agtgtgacaccttggcttgactgttggaagacagaaagggcatgtagtttataaaatcagccaaggggaaa
 atgcttgtcaaaatgtattgtcgggtattttgattaatagtttatgtggcttcattaattcagagttactct
 ccaatatgtttatctgccccttcttgtctgataatggtgaaaacttgtgtgatgcattgtatatttgattta
 ggggtgaactggatgtcttctgttttcttctttag**TGCAATTACGTTGTCCCTGCCACACTGGTCATTATCAT**
CTTCATCTGCTTCCAGCAGAAGTCTTATGTGTCTCCACCAATCTGCCTGTGCTAGCCCTTCTACTTTTGCT
GTATGGGtaagtcacctctgagtgagggagctgcacagtgaggaatggcatttggtgcccagtggtcagaagga
 gggcagggactctcagtagacacttatcttttgtgtctcaacag**GTGGTCAATCACACCTCTCATGTACCC**
AGCCTCGCTTTGTGTTCAAGATCCCCAGCACAGCCTATGTGGTGCTCACCAGCGTGAACCTCTTCATTGGCAT
TAATGGCAGCGTGGCCACCTTTGTGCTGGAGCTGTTTACCGACAATgtgagtcagtcagagagaacactcct
 gctgggatgagcatctctgggagccagaggacagtggttaattgtgatcttattccacttgtcagtggtatt
 gacactgctgactgccttgcctgtcttcagagtctgtcttccctgagaaggcaaagcaccttcttctctg
 ctgtgccttacattttgtgggtcaagccttccagtttcttttgacagtttttttacttcttcttttttca
 atgttgctcttaccagagtagctcctctgccttccactttacacatgagagctgggcgacgzcattcagtc
 ctaaggcttttaccatcacctctcttgggtgttttattgtcatctctaagatcaatgcctttagccttgatc
 ataacttgaactctaattctcaaattctcacttgcctagtggtattgctccatttagatagtatatagatacc
 ccaactggatatgtccttagttttcttctcccttggaaacttaatgcttttcttgccatccctgtcacactca
 gtggcactaccatccactcgggttggccaagctggctcttagagttatccttagatgcttgctttgctgttgca
 gatttcccacattcaactgggtatgttgtcagttcttccaggtatggacctctaaaataaggcttccctctcc
 attccggttgtcattgccttctgccaacacagcacacaaggccttttacagttgcacaactcttccctgtcc
 ataccacacacaccccttccagctgtaagcttcagatgagttgcctccaaccaccatgctcctgtaggcct
 ggcttgaaattgcccttctctgtcacaggggtcgtgttagtatatcccttgccctcaagatttagctaaaatg
 tgaagcttcccttacctgctgggaggtgtctctcttcttctctgtgtctcagagtccttagtccatgcctc
 cagtacaacgtacatccacttacatggtaatttccctgtttacatacttttccctactcggagtgaggctgtgt
 tcttaataattttgcctctcccatgccttagcacagtgcatccagcgtatagcccttattcagttggtaga
 tatttggccactgttgcttgtgggatcataagttctgatgtatttgagaagaatttctaaaattctgacaa
 aatcctgaaactcaaatattgaccacagacatgagcaatttgcttttcaaatgctaagggtttttaatggat
 ttgctttaattaaatctagcctgtttctaaagctttattcattatttctccatactcagagcatttctccaga
 ttttctaaagaatagaattttattgtctacatatcatcagctatgcctgctgctatttaattgggtatctgaat
 taaaaggctcgtgttctgcttagagaatcaaattttttcttccactcccatatttccagaacttgatacatttt
 taggataaaccatgaatgacacccgtttcttctccctcaccctcccttccctccctttttttttttttttt
 tttttag**AAGCTGAATAATATCAATGATATCCTGAAGTCCGTGTTCTTGATCTTCCCACATTTTGCCTGGG**
ACGAGGGCTCATCGACATGGTGAAAAACAGGCAATGGCTGATGCCCTGGAAAGGTTTgtgagtgaaagcag
 tggctgtaggatgctttaatggagatggcactctgcataggccttggtaccctgaactttggttttggaaga
 agcaggtgactaagcacaggatgttccccacccccatgccagtgacagggctcatgccaacacagctgggt
 tgtggcatgggttttgtgacacaaccatttgtctgtgtctctgatagcattgagaaaagtgaagggcagtt
 ttgaaggtaaggaaaatagtgttatttgccttgatccactggctcatgccactgtctgggttgggttagaagc
 actggaaaagtcaaacataactttgagaattaggtgatcaggggaatcagaaggaaagatgcaaaccttggc
 tcttttaggcgaatcatgtgcctgcagatgaggtcatttattatcttttacacagtcataaaaattataatg
 tattacatcttttctacctttagaatgggttaaaaaatatttctccggtagccatatgattatttcatcca
 ttagataatatagtcaaatgggccatgttatttactgttcatagaagaggggctttttgcaacttgggctac
 aaaggagatatgtaagggaatttaagggaatgggtacatggaactagatttaattgaatctagtgggttaattg
 attcactaggatatatgctactgaaaggggaatctgcttaaagtgccttctgatatttatttactaaaac
 ttagaattttatataaaaactgactgtgaaaattacttgggtcggttgccctttttaaaggatttttggcat
 gtctcattaaaaaaaagaaatactagatatcttcagtgaagttacaaatcgaatacacattggctctgaaatt
 ctgattgatactgggtcataaaaagttttcccaaatcagacttggaaagtgatcactctcttgttactcttt
 tttccttgtcatgggtgatagccatttgtgtttatttgaagatcgggtgaatttttaagggaacataggcccaa
 tttgaggaagggccatgggttttgatccctccattctgaccggatctctgcattgtgtctactag**GGGAGAA**
TCGCTTTGTGTACCATTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCCATGGCCGTGGAAGGGGTGGT
GTTCTTCCTCATTACTGTTCTGATCCAGTACAGATTCTTCATCAGGCCAGgtgagctttttcttgaaccc
 gtggagcacctgggtgaggggtcacagaggaggcgacagggaaacactaccaatgggggttgcatgaaact

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

gaactcaaaatgatgtgataaaactgatttttctgatgtgggcatccgcagccccctccctgcccacctctgg
agactgtggcaagtaggttttataatactacgttagagactgaatctttgtcctgaaaaatagtttgaaagg
ttcatttttcttgttttttcccccaag**ACCTGTAAATGCAAAGCTATCTCTCTGAATGATGAAGATGAAGA**
TGTGAGGCGGGAAAGACAGAGAATTCTTGATGGTGGAGGCCAGAATGACATCTTAGAAATCAAGGAGTTGAC
GAAGgtgagagagtacaggttacaatagctcatcttcagtttttttcagctttatgtgctgtaaccagcag
tttgcctgacttgcttaataaaaagggcatgtgttcccaaatgtacatctataccaaggttctgtcaatttta
ttttaaaaacaccatggagacttcttaagaattcttactgagaattcttttgtgatatgaattcccattct
cgaatacatttggtttttatatgcttacattttatgtgttagttattaaaacataactaatattgtatatctagtc
aaaactgaggttagagagaataaatgggtgattttgagtttgagtttcatagtcctaaaagctgatataattgc
ctgtgttcaagagggctatatacagccctctagatgccagcatctccaaattttacttttttgggaatctgta
cagtatttgaataatttttattacaaatttctactctgtggaatttaatttttaaaatacctgcaatacata
tatatgttgaatagatgaaaaattatgtagatrataatgaatgatacgggttctaaaaagacaggttaaaaag
taagttcactttttattttgagcttcagaatcattcagaagccagtcgccacaaacgcagaccaaggtcttg
gcacatcaaatatgcctatggcttagggttattgacaagtcttatgttgcagtgtatgtggtttatagtcct
gccttccacagttgcttgggagagctgtgagtcactgaggtctatgaatgtttacattttgtttgttgcag**A**
TATATAGAAGGAAGCGGAAGCCTGCTGTTGACAGGATTTGCGTGGGCATTCTCTCTGGTGAGgtaaagacac
tttgcctatatattgcgtttgtccctattagttcagactatctctaccaatcaagcaacgatgctcgttaaga
ggtaaaagtggatttttaagggtctctgtattttatgccaggatggagcaattagtcacgagaagagagggac
cctgtatgtcaagagaatgatttcagagaatccaatacaatttaagaaaaagcatggggctgggagcagtg
ttcactcctgtaatcccagcactttgggagggcaggtgggaggactcacgaggtcaggagattgagaccat
cctggccaacatggtgaaaccccactctctactataaatacaaaaattagctgggcatagtagtgcatctctg
tagtcccagctactcgggaggtgagggcaggagaattgcttgaacctaggagggggaggttgcccagattgc
gctgctgcactccagcctggtgacagagtgcagactcatgtcaacaacaaaaacagaaaaagcacgcacatct
aaaacatgcttttgtgatccatttgggatggtgatgacattcaaatagttttttaaaaatagattttctcct
ttctgggttccggtttgtgttcttttatgcccttttgccagagtaggtggtgczaatttggctazgctggctt
tcattactgtttttcaczaczaattaacztttggcctcaacttgacaactcaataaatatttataaatacagc
cacacttgaattgggtccattatgaaatacatatttaaatatctatacagatgtgtttaaaccagaaaaatat
ttgattcttctctgatatttaagaattgaaggtttgaggtagttacgtgttaggggcattttatattcatgtt
tttagagtttgcctatacaacttaattctttccttttcag**TGCTTTGGGCTCCTGGGAGTTAATGGGGCTGGA**
AAATCATCAACTTTCAAGATGTTAACAGGAGATACCAGTGTACCAGAGGAGATGCTTTCTTAACAAAAAT
AGgtgagaaaaagaagtggcttgatatttgcctgcaaaagactttgttttttaattttatttaagaaataggttgt
tatttttgattacagtgggtatttttagagttcataaaaatgttgaaatatagtaaaagggtaagaagcacat
aaaatcatccatgatttcaatatctagagataatcacaatttacatttcccttcagtcctcattctctctt
taacagcttttattcaggtataatttacatacaatataatttgcttggtttttaagagtataatttagtgatt
tttggtaaatgagagttttgcaaccatcaccacaatccagttttagaactttccatcacccacatctgt
cttatatacacatataaatgtgccatacaattgagatcatactgtatgtagaatttaaaattagtttttatt
gttaatgagtgattatgaatatttcccagtggttacatttccctaagatgtggaattttacattgctacat
aaaatccccctatgtacatgtacctataattttatttaataaattccttataaatgttggacacattagtttc
catttttcaactatgtaaatatgtccctgtatacatcttttattatttccctcaggaacaattcctacaaagta
aattgccctctctaaagagcatataaaattgactgagccaccggttagggcattttctgagactgcacaggtca
caaagcaatctgatctttgggaatacagctacattttataggtctcttagataatgttactcttaagtacttt
aaatatgtggggctctctgggcttttttttttttgagacggagtttactcttactgccagggtggagag
caatggcgcgaccttggtcactgcaacctccgctcccaggttcaagcgattctcctgcctcagcctcctg
agtagctgagattacaggtgccgcgccacaatgcctgcctaattttttgtattttcagtagagatggggttt
caccatgttggccagactggtctcgagctcctgacctcaggtgatccacctgcctcagcctcccaagttct
gggattacagggcatgagccactgcgcccggcttctctggacttattatgtggagagatagtaaggcagtg
gctttcagagttttttgaccatgacctgttgggaaatacattttatatctcaacctagtatgtacacacag
acatgtagacacatgtataacctaaagtttcataaaagcagtaacctactgttactaattgtagtgcactctgc
tatttcttattctaccttatactgcgtcattaaaaaagtgtggtcatgacctactaaatttatttcccaaa
ccactaatgaacaatgactcacaatttgaacacactggacagggggatagccaataaaattgaaaagagcaa
ggaaattaatgtattcatgatctcctctcctgtctcttacatttttgagtagcaatgtaaaggaatcctaa
gagaacagacattctgggaatagcaggcctagcgtgcacaactgctttcctaggttgctcctagtaacca
gctcctgacgcataatagcagtggcagtaataaccagcccataagtaaggtttgtcacagggactggttgaag
aactgatttgggtgggtatagctgtgagggcctggcaggtgtccacgtgtgctcaatcctaattctgaaaa
aggctgacctgggggtgctaattagatacacagagaggaatgaatgctgccagaaggccaagttcatggca
atgccgctgtgggtgaggtgcagtcacgtctggaacgtgaacactgaacttctcactatgtgatctctc
acttgactggcttcatagaaccccaagccaccccccacacataaattgtgtctctaggttctgtgttgc

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tcacactcaaaatttctctgggccttctcatttgggtgcatgtgaatgggtgcatatgagtgaaagtctaggatggg
gccttagcggttaaagccctggggtagtgtagctgagattggttggttaaagaatgtgcagtggttggcatgacc
tcagaaattctgaaatgggactgcacctgcagactgaagtgttcagagagccagggaggtgcaaggactggg
gagggtagaggcaggaaccctgcctgccaggaagagctagcatcctgggggcagaaaggctgtgctttcaag
tagcagcagatgtattgggtatcttggtaatggagaagcatactttacaggaacattaggccagattgtctaa
ccagagtatctctacctgcttaaaatctaagtagttttcttgcctttgcagTATCTTATCAAACATCCATG
AAGTACATCAGAACATGGGCTACTGCCCTCAGTTTGATGCCATCACAGAGCTGTTGACTGGGAGAGAACACG
TGGAGTTCTTTGCCCTTTTGAGAGGAGTCCCAGAGAAAGAAGTTGGCAAGgtactgtgggcacctgaaagcc
agcctgtctcctttggcatcctgacaatatataccttatggcctttccacacgcattgacttcaggctgttt
ttcctcatgaatgcagcagcacaaaatgctgggtcttctgtatctgctttcagggtggaaacctgtaacgggtg
gtggggcagggctgggtgggcagagagggtgctgctcccaccacacgagtccttctcctgctttggct
cctcaccagttgtcagggttatgattatagaatctagtcctactcagtgaagaactttcatacatgtatgtg
taggcacagcatgataaaattcccaagccagaccaaagtcaagggtgctttttatcactgtagGTTGGTGAAGTG
GGCGATTTCGAAACTGGGCCTCGTGAAGTATGGAGAAAAATATGCTGGTAACTATAGTGGAGGCAACAAACG
CAAGCTCTCTACAGCCATGGCTTTGATCGGCGGGCCTCCTGTGGTGTCTCTGgtgagataactgtggatgg
aaaactgttgttctggcctgagtggaacacatgactgttcaaaagtcctatatgtccagggtgtgtgtatga
ttggctgtgtctccccagggacagcagagcaaccttggaaaagcagagggaagcttctccttggcacaca
ctgggggtggctgtaccatgcctgcagatgctcccaaatagaggcactccaagcactttgtttcttagcgtga
ttgagggtggatattgtgatttgatcttctctggaacattctttctaatacatctttgtgttcatccctgaa
aatgaagagtgtggacacagctttaaaatcccaaggtagcaactaggctcatagttcctztacacacgggata
gatgaaaaacagatcagactgggaagtggcccttgacctttttcttctgtagataagagcattgatgttat
tacgggaagaagcctttgaggcttttatgtattccacctcgggtctggaatttggtttctgtaaggctaacagt
tgcaatatactagggtaatctgagtgagctggaattaaaaaaaaaagggaatttcaccccaatcttatactg
acttcaatagagggttccagacaaaaagtgttttggatatacttatactcagtcagtaaaagataattacaacta
aatggcctttttctcctcctattttatttggagaatttaattacataaaaaagctactcagaatatttgagtt
tctgcatcaataagacatttataataatgacctgtttacaaatgaatttgaaagttactctaattctttg
attcatcaagaaataactagaatggcaagttaaaatttaagctgtttcaaagatgcttctgcatttaaaaac
aaatttatctttgattttttttccccccagcaataagacttattttatttctaattacagGATGAACCCACC
ACAGGCATGGATCCCAAAGCCCGCGGTTCTTGTGGAATTGTGCCCTAAGTGTGTCAAGGAGGGGAGATCA
GTAGTGCTTACATCTCATAGgtccgtagtaaagctcttgggttctcactgtgggatgttttaactttccaag
tagaatatgcgatcattttgtaaaaattagaaaatacagaaaagcaagagtaaaacaattattacctgaaa
ttatatatgcataattcttcaaaaaatgcaagcccagttataaatactgctctttttcacttaatatattgtaa
acattattccaagtcagtgcatttagggtgtcatttcttataagctggatagttatccattaggatatactctt
atttaactattccccctttttagacatttggattatttccaacttggttcacaattgtaaacaccactacac
tgaacagcatcatccctatatccacatgtacttgaacagaatacaattccctaggaagctggaatgctgga
agtcatgggtgatgttctcatgggttacagagaatctctctaaaaactaaaacctctttctgttttaccgcagTA
TGGAAGAATGTGAAGCTCTTTGCACTAGGATGGCAATCATGGTCAATGGAAGGTTTCAGGTGCCTTGGCAGTG
TCCAGCATCTAAAAATAGgtaataaaagataatttcttgggatagtgcctagtgagaaggcttgatattta
ttcttttgtgagtatataaatgggtgcctctaaaaataaagggaataaaaactgagcaaaacagtatagtggaa
agaatgaggggtctgaagtcgaactgcattcaaatctgtctttaccatttactgggtctgtgactcttgg
gcaagttacttaactactgtaagagttagtttccctggagatctacctcctagctttgtgctatagatgaa
atgaaaaaaatttacatgtgccagttactgggtgagagcgcaagctttggagtcaaacacaaatgggtttgcat
cctggccctaccaattatgagctctgagccatgggcaagtgaactcctgggcctcagtttctctgttaa
catctgtcagacttcatgggtccagggtgaggattaaaggagatcatgtatttacagcacatggcatgggtgct
tcacataaaaataagttatttagttaaataagataactgggttcttctctcagaaacttatttctgggcctgccagg
ggccgcctcttttcatggcacaaagtgggttcccagggttcagttattcttttaaatagttttctggagatcc
tccatttgggtattttttcctgctttcagGTTTGGAGATGGTTATACAATAGTTGTACGAATAGCAGGGTCC
AACCCGGACCTGAAGCCTGTCCAGGATTTCTTTGGACTTGCATTTCTGGAAGTGTCTAAAAGAGAAACAC
CGGAACATGCTACAATACCAGCTTCCATCTTCATTATCTTCTCTGGCCAGGATATTTCAGCATCCTCTCCAG
AGCAAAAAGCGACTCCACATAGAAGACTACTCTGTTTCTCAGACAACACTTGACCAAgttaagctttgagtg
caaaacagatttacttctcagggtgtggattcctgccccgcacctcccgcccataggtccaagagcagtttg
tatcttgaattgggtgcttgaattcctgatctactattcctagctatgctttttactaaacctctctgaacct
gaaaaggagatgatgcctatgtactctataggattattgtgagaatttactgtaataataaccataaaaaac
taccatttagtgagcacctaccatgggcccaggcatttttacttgggtgcctaactcctattttaaattagataaaa
aagtaaccaaataggctcctgacacttaagaagtactcagtaaatattttctcctccttccctttaaataaga
ccgtatgtgccaaagttaaattggatgactgagcaggttgggtgatgtaggggtggggggcgatatagaaagtcag
tttttggccgggcgtgggtgggtcatgcctgtaatcccagcactttgggagggtgaggagcaggcagatcatg

bioRxiv preprint doi: <https://doi.org/10.1101/144444>; this version posted April 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

aggctcaggagatccagataatcctggccaacaggggtgaaaccccgctctctactaaaaatatacaaaaattagct
gggcatgggtgggtgcgcacttgtagtcccagctacttgcgagggtgaggcaggagaattgctcgaacccagga
gggtggaggttacagtgagccaaggtctcgccactgcactccagcctggggacagagcaagacccccatttcaa
ggggggaaaaaaagtcctatttttaagttggtattgcttttttcaagtatcttccctccttcacacacagtt
ttctagttaatccatttatgtaattctgtatgctcctacttgacctaatttcaacatctggaaaaatagaac
tagaataaagaatgagcaagttgagtggtatttataaaaggtccatcttaatcttttaacagGTATTTGTGAA
CTTTGCCAAGGACCAAAGTGATGATGACCACTTAAAAGACCTCTCATTACACAAAAACCAGACAGTAGTGGA
CGTTGCAGTTCTCACATCTTTTCTACAGGATGAGAAAGTGAAGAAAGCTATGTATGAAGAATCCTGTTTCTAT
ACGGGGTGGCTGAAAGTAAAGAGGAACTAGACTTTTCTTTGCACCATGTGAAGTGTTGTGGAGAAAAGAGCC
AGAAGTTGATGTGGGAAGAAGTAAACTGGATACTGTACTGATACTATTCAATGCAATGCAATTCAATGCAAT
GAAAACAAAATTCATTACAGGGGCAGTGCCTTTGTAGCCTATGTCTTGTATGGCTCTCAAGTGAAAGACTT
GAATTTAGTTTTTTTACCTATACCTATGTGAACTCTATTATGGAACCAATGGACATATGGGTTTGAACCTCA
CACTTTTTTTTTTTTTTTTTTGTTCCTGTGTATTCTCATTGGGGTTGCAACAATAATTCAATCAAGTAATCATGG
CCAGCGATTATTGATCAAAATCAAAAGGTAATGCACATCCTCATTACTAAGCCATGCCATGCCAGGAGAC
TGGTTTTCCCGGTGACACATCCATTGCTGGCAATGAGTGTGCCAGAGTTATTAGTGCCAAGTTTTTTCAGAAAG
TTTGAAGCACCATGGTGTGTCTGCTCACTTTTGTGAAAGCTGCTCTGCTCAGAGTCTATCAACATTGAATA
TCAGTTGACAGAATGGTGCCATGCGTGGCTAACATCCTGCTTTGATTCCCTCTGATAAGCTGTTCTGGTGGC
AGTAAATGCAACAAAAATGTGGGTGTCTCCAGGCACGGGAACTTGGTTCCATTGTTATATTGTCTCTATGC
TTTCAGCCATGGGTCTACAGGGTCTATCCTTATGAGACTCTTAAATATACTTAGATCCTGGTAAGAGGCAAG
AATCAACAGCCAAACTGCTGGGGCTGCAAGCTGCTGAAGCCAGGGCATGGGATTAAAGAGATTGTGCGTTCA
AACCTAGGGAAGCCTGTGCCCATTTGTCTGACTGTCTGCTAACATGGTACACTGCATCTCAAGATGTTTAT
CTGACACAAGTGTATTATTTCTGGCTTTTTGAATTAATCTAGAAAATGAAAAGATGGAGTTGTATTTTGACA
AAAATGTTTGTACTTTTTAATGTTATTGGAATTTTAAAGTTCTATCAGTGACTTCTGAATCCTTAGAATGGC
CTCTTTGTGATCAACCTGTGGTATAGAGGAGTATGGCCACTGCCCACTATTTTTTATTCTTATGTAAGTTT
GCATATCAGTCATGACTAGTGCCTAGAAAAGCAATGTGATGGTCAAGATCTCATGACAZTTATATTTGAGTTT
CTTTGAGATCATTAGGATACTCTTAATCTCACTTCATCAATCAAATATTTTTTTGAGTGTATGCTGTAGCTG
AAAGAGTATGTACGTACGTATAAGACTAGAGAGATATTAAGTCTCAGTACACTTCTGTGCCATGTTATTCA
GCTCACTGGTTTACAAATATAGGTTGTCTTGTGGTTGTAGGAGCCCACTGTAACAATACTGGGCAGCCTTTT
TTTTTTTTTTTTTTAATTGCAACAATGCAAAAGCCAAGAAAGTATAAGGTCACAAGTCTAAACAATGAATTC
zTTCAACAGGGAAAAACAGCTAGCTTGAAAACCTTGCTGAAAAACACAACCTTGTGTTTATGGCATTAGTACCT
TCAATAATTGGCTTTGCAGATATTGGATACCCCATTAAATCTGACAGTCTCAAATTTTTTCTCTCTTCAAT
CACTAGTCAAGAAAAATATAAAAAACAATACTTCCATATGGAGCATTTTTTTCAGAGTTTTCTAACCAGT
CTTATTTTTCTAGTCAGTAAACATTTGTAAAAATACTGTTTCACTAATACTTACTGTTAACTGTCTTGAGAG
AAAAGAAAAATATGAGAGAACTATTGTTTGGGGAAGTTCAAGTGATCTTTCAATATCATTACTAACTTCTTC
CACTTTTTCCAGAATTTGAATATTAACGCTAAAGGTGTAAGACTTCAGATTTCAAATTAATCTTTCTATATT
TTTTAAATTTACAGAATATTATATAACCCACTGCTGAAAAAGAAAAAATGATTGTTTTAGAAGTTAAAGTC
AATATTGATTTTAAATATAAGTAATGAAGGCATATTTCCAATAACTAGTGATATGGCATCGTTGCATTTTAC
AGTATCTTCAAAAATACAGAATTTATAGAATAATTTCTCCTCATTTAATATTTTTCAAATCAAGTTTATGG
TTCTCTCATTTTACTAAAATCGTATTCTAATTTCTTATTATAGTAAATCTATGAGCAACTCTTACTTCTCGGT
TCTCTGATTTCAAGGCCATATTTTAAAAAATCAAAAGGCAGTGTGAACATTTTTGAAGAAAAACAACATT
TTAATACAGATTGAAAGGACCTCTTCTGAAGCTAGAAAACAATCTATAGTTATACATCTTCATTAATACTGTG
TTACCTTTTTAAATAGTAATTTTTTACATTTTCTGTGTAAACCTAATTGTGGTAGAAATTTTTTACCAACTC
TATACTCAATCAAGCAAAATTTCTGTATATTCCCTGTGGAATGTACCTATGTGAGTTTCAAGAAATTTCTCAA
ATACGTGTTCAAAAATTTCTGCTTTTGCATCTTTGGGACACCTCAGAAAACCTATTAAACAACCTGTGAATATG
AGAAATACAGAAGAAAAATAATAAGCCCTCTATACATAAATGCCAGCACAAATTCATTGTTAAAAAACAACCA
AACCTCACACTACTGTATTTTATTATCTGTACTGAAAGCAAATGCTTTGTGACTATTAAATGTTGCACATCA
TTCATTCACTGTATAGTAATCATTGACTAAAGCCATTTGTCTGTGTTTTCTTCTTGTGGTTGTATATATCAG
GTAAATATTTTTCAAAGAGCCATGTGTCTGTAATACTGAACCACTTTGATATTGAGACATTAATTTGTAC
CCTTGTTATTATCTACTAGTAATAATGTAATACTGTAGAAATATTGCTCTAATTTCTTTTCAAATTTGTTGCA
TCCCCCTTAGAATGTTTTCTATTTCCATAAGGATTTAGGTATGCTATTATCCCTTCTTATACCCTAAGATGAA
GCTGTTTGTGTCTCTTTGTTTCATCATTGGCCCTCATTCCAAGCACTTTACGCTGTCTGTAATGSGATCTAT
TTTTGCACTGGAATATCTGAGAAATTGMMAACTAGACAAAAGTTTCAACAACAGATTTCTAAGTTAAATCAT
TTTCATWAAARRRAAARAAAAAATTTTGTATGTCAATACTTTATATGAAGTATTAAAAATGCAATTT
TCTATGTTGTAATATAATGAGTCACAAAATAAAGCTGTGACAGTTCTGTTGGTCTACAGAAAttacttttg
tgcatttgtggcaccacactggttgaaggggtataaagccattagaaaagtagaggggaagtgatttggat
caaaaggaaaaactttagaaaagacttcaatgttcccttaatacataaaagagaactgaggggactacttgaa
aataaaagggttgttttgcattttcatggttgggttaagatactgagztaacwgggtattaagtgttagaggtttt

tagataaatattctgcttaattgattatgaagctgcactgagatttctgaaaatgctctgtagctgagcttat
ttaataaatgttcacttgggtataggggaagctacaaaggcagccttcagtgctcttttgtttattcaaccaa
aaatataaggacacaaatgtagcagttatactgggaaggtgctgggggtgggtggcaatggtagcaggaaggc
gaagtagatatggaaacagaaatgataactaatatcggtgattccttccttttttctgttrataagtgtgtg
cagacaacatatgagcagtgctgataaatgtaaatgtattkttcatagctcattaagaatcagtttcagaaa
gagatgtctgcttattttkgctrcttgaagaatccctgtcaaacagtccttttsaggaagtacaagaggctgt
ctctattttgtgacctcaggaatggctgtgacagtgctgtgagcagtccttttctgtggcacagatctgaac
tttgtgtgcagaaaaatccttggcttcaagtgaagcaagatgccccctgagcatcagcatcacaacttcatcc
tctatcttgaagttcatgttatagtgaactttaatgaaatcatagaacactgtttcttcgtgzaacaatgac
gagggagaggaaaaaactttattgaaaaataaaaaggcaggttaatttagatgaaaatatgttaccatgagg
ttttgtttttgctttttgtttttgtttttgagaacaagaatctcgctctgtcgtccaggctggagtgcagcg
gcatgatcttggctcactgcaacctccgctccgggttcaagcgattctcctcagcttcccaagtagctgg
tactacaggcatgcgcccaccacaaccagctaattttttgtatttttagtagagatgggggttctactatagctt
ggccaggctgggtctcaaactcctgacctaaaggtgatccttctgccttgggctcccaaagtgctgggattaca
ggcatgagccaccttgcctggccctacctatgagccttgactaaaacattcttctatctgtagaaaagccca
aaagaacttttccagattcaaaaaacttggcactttgtaatggtaattgtttacattaaagtaaaaaaaaaa
aaaacccacttagcttcagttttcaagtgtttactgtgtgtcatgcacttcatttaattctcaacacctgc
cctatgaggtaaaaagtagcatattttacatatgagtaaaattacagctcagtgagataagaaactcgtccaaagg
tacaggttcagtcagtgagggaggggttctttttgttgaagttaggtatcagttaaaaattgaccttgtaaaa
tcacatcagcatcaatatatacatataatttaacaaatattttattgaactttactgtatgccagatacttctcta
ggtagtaggggttacaatgtagaagaaaatagaattcctgctctaataatgaatttatatttcagtggtgaaaga
tgatgtgtggacaaacacatctaattgtattttgacagcaatgggtgctaagaagaaaataagacatggtaat
tgagatacaatggagagaagtcagagatggccttttctgaggagtgcagatgctcagaatcgaataaaaagag
caggagcagcctttcaaggtgggtgcgaaggatattcccgaagaaagaataagtggtgcaaaaagcccgagt
gggaaaaagccttggaagctgaggtggtaaaagagctatctgtactaatgtctcttctgtactgagttagcaag
gacaagagtggcaggggggtggaactggagagatagcagggaccacatctcacaggatctcgccagcctttt
aaggtattttggatttttattgttagtgcaacaaggagcactggagagttttaaacagtagtggtgtgacctta
ctgtttcagaagaacactggctactatgggaagaaaggacagtaggaagactaatagataatgggtggattg
aactaaagtggtagcaacagatagggagctatggtaattgttcagtatccactttggagatagatccagcagg
acttgctgacagactgcagtgtaagggctgagggaaagagagctagcaaggtgacttctagtttgtgacctga
actaggtagatgggtgtgttaaaaatgcagtggttatttcagctctgaagctcaggggcaatgatggaaaata
agataaatgtgtgcaagttggctgggctgggtgggtgcacgcctataatcacagcactttgggagggccgagac
gggtggatcacttgaggtcaggcgtttgcgaccagcctaacatgggtgaaaccccatctctaataaaatacaaa
aaaattagctgggcatgggtgggtgcagtcctgtaatccaagctacttgggaggctgagacaagagaatcgctt
gcacctgggagggcggaggtttcagtgagctgagatcgaccattgcactccagcctgggcaacaagagcgaa
actccatctaaaaaaaagtggtgcaagttattgggtacgtatatatttaagcctcgggactgaatgagatta
cctagagagaatgaagataggggaacattcttgagcctgagatcctataacatttagaagagaagctagccag
ggaaattgagaggggactggccagtgagagagagtaaaatctgaagagtataatgccatagtgtgaagctcag
agaattattgtttcaggaaggcagaagttgtcaattatgtccaacattgctgagaagatgagtaaggaatgt
aatgagtaaaaggcaggaatggctattgaatttgttaagacagaggctccttgggtgaccttccctaaaggtcat
ttcagcatgacaaggatgcagacctaaactggaatagattttcaggggaagactgagaggaaggaaggggaga
ctgcaaatatatgcaaatctatatattttgtcttgaagggaacagaaaaacaggacaacagctgggttacaga
acttttatttggattcttctccttgtctgttataataatataaggctgcacctgggcatcatcgtgcccttctctt
ttgattctgctgttttgccttttaactcctacaggcatttcttctggcaggagccttgtacacactctgggtgg
ttcatgaccaccaggacttgcctttcaaggcagaagcctgtgattcaaaaacttccattccttccaatcg
actattataaaaactcattttgtctagtgtccatactcctttaccttaggatcctgggccaactgtcactaaa
gaaaacaaaccatttctcctcctgggggaatgtgcttctttcaagaagggtgcttagagaatagacattcta
ggtagcagcagttatgaaagaggtgtctctaatagaaccttcttccaacatactccaggttcacctt
cctttgaagacagaaagtactacaacttcagttgccaccctagtcaggttcataggatttactcagtaggtt
aagagatgggtgctgagaatgaagtcagagtttgagtctacaagtggtcatttgcttctcctttgggacttct
gtggcaggggtgtggaaccaaactcagattgggtgcatactctgtttctccttctgaaaatacaattcaaatacc
ttcatagctctgtgacagcaacttataaccaacagacaatggattattttttattttatttgatggcactcccc
atagccagaacttctactgagtcattcatcgtgtattccaggaatacttttccaaatggagtagtgaggcact
tccaataacagttatgtaactgttaacagtggtcctcaaatgacatgtgaaggaatgatgtacttcttttt
cagactcacagatctcatatgttaattcagtcacaagtagttagggcttatcacatgccaggcacaaggttaa
tatgacaataatgaaaacaaaaacccttgctgtcatagaattcggtttctagtgggaagagacatacatataa
caaatcatatataagtaaaacattaattatgggtgataacagaaaaagaaagcaggggagagggatagtgtct

tagttcaggttgctgtaacaagagtaccatagacaggggtggcttamacaacaaacatttctttctcacagat
ctagtggctgggaagtctaagaccaggggtgccagcatgaagaagttctgggtgaaggcttacattctggtttc
ctcacagtgagggctcttcacgcccttataaatgcactaattctattaccacatgacctcatctaaaacta
actcacaaggccaaacctcctatgttgagattaaggcttcaacataggaattttggggggacaaaaacatt
cagttctctagcagatagggggccactgtggggaggggattacaatcttaaatagggttgtcagggaaggcttca
tgaggggcaaagacctgaggagggcatggaaatgatccacgcagatgcatgtctgggggaagagcattccaggga
gaggcagtgaggaaatggaccctcaggttaagagcatgttttagcaagttcagagaaaagcctcttaggtcaggtc
tggagcgagagtggaatatgtaacaggacagatcatgtaaagacttgtagaccacagtttagaggcttggaca
ataactctgaggatgacatgaaggtattgggaggggggtgttgagcagaggagtgatgatctggtttatat
tttagattattctgactgctatgttgaaaaataggggttgagagcagaggggaagggcacaccagtttaaga
ccacttcccaaaaaatccagatgatgatgggcttgaaccaggagagttgctatagggatgtgaagcaatcaaa
ttctagatatttgggacctgctgacaggtacataaatgggtgtatgagagaaaaggcagtggtgagaatgactct
gaagttctgtggcctgagcaaaacagaaacctggagttgccaatagctgagataaagtttctgggttaagatc
aagagctcagttttggacaagaggggtgtgagatttctaataagacattcaggtggaggtgtggggaacacac
ctggatgtgatgagcctggaattcaaatccacatgtgtgaatttgaaaaagattcagaaaagaggtgtgga
ctggaggagctatccacacacagatagcatttgaagccacaagactgggggtgatgacgcagggagtaagta
taaatagaaaagagaagtggtccaaggactcagtccttacatcagaggtcaaagagacaagatggagactag
gagatatgaaggcttatgcagtgagtggttcggagaaaacgaagaaacttccagaagaaaggagaggtcaac
tgctaacaagtcaagtgaatgagaactgggaaatgagttgacaagagcaatttggctggcgtgatcagaga
atgtttaagagataatgtgaggggaagaaaatggaggcaggaagtatgaactactcttttgagtcataattgta
aaggaaaaaaatgggggtgattttactcttttttttttttaagataggaaaaaaagtatatatttatatgctg
aggagaagaatccagttgaaggggaataactgtggcttcaggagacaggaaaaactgcctaagccacatccttc
agcaagtgagtggtcaagtgagggtttgtctttgcttggggcacgggcttcatccacagtgaaggaga
gaaggcagggcatgtgggttggctgtagaaagcgggcagacgggggtgggaacttatgaagactcttctgat
tgttattttctctgttaagtaaaaggaaaggtcattggctaagaatgaagatggacgcgcaggtgctgggag
tcatcgaagagaatgaaagaatgaataaaaatagagaaaaatacaatgctattatataaacctatacatttaa
attactttcctcttcaaattagacccccctggaggcaaatctctgtcatgttaagtttaggaaaagtcata
atcacttggagctaaagtatgaagaataaggtatatgatttaaaccacataattctatttttgcctaaaccg
gggaagcctttaaggtaacgtagagtgtaacctataaaaaagataaaaaatgtcagcttcccccttcccttaa
atatagtaaggctcaaaggtaattgatgttcataaccctaagggtactataacagctctgaagcaaaatattaa
atgtcttctgtagtattttagttaacttttaattttattggtaatttatgtgtagttaaccacataacttggtg
cctcaactatttcacataaatgtgagatgtcttaggcctctttgaactctgcatggaaaagaatgctgagga
gggtatgtctgtttaagctgccactgctgctctcataaataatatctcacattgcataatagcccagataa
actgcagatcgcttagagcctgataaatgagaggaaactgtacagtttggtttccaaattgcccctggacag
ctaattcacaatttactattttatagctgtaagattttaaaaatataaaaataacatagctgtaaatacagat
tatcagaagatagttctgtaaaaatgtaaatataatataagctgtaaatatacactatcggaagtaccaaactc
agtttttagtaggttagagggcatcaagtaaaatggaaatgaaagatgtttcaaggaaaactccaggacatctgt
ggcagtaactaaagaaaaccttcccttctcaggttcccagcatgctattttattgatgtaacaacattttcaata
aagttggtaaatatcactatattactgtcttatagcaacatagcaagagcttttagagatcatataagttata
aaatgtgaattttaaaaaacaatgaatatgcaggatttttattagggaagcgtttccataaccataaatat
ttctttaaaacaaataaatgtcccaagatctctgttagtgatccaaactaagtagaatttagtaaaattaat
tataaatgaacaatttcagcatataaaaccaacaagtccttttctagatttttaacactgtgaccaattgcat
tattttccaagttagaatgactaataatcaatgaatgtaaaagcaataattaatacagatgacattgtactt
ttccacagtaaaagaaataaacaatctaataatttttataaatccatttttatatcacaaaaataacctttacta
agcaaatttttttaaatctcagggaactatagacatgatgaaaagatagatattttatataaataaattcaaa
aatactgtcagggaaggaaatgtaaaatccttatttgagtataaagaaaatgctataaagcaatgagttatca
aaatacagaagaggtatttctaaaacaaatgaaaaaccaagatgatgaaatagtgaactacttctaattgtg
taacagatactgaaatgccagggtgaaagtgaactgaattatttcttaagcagtgaggagaatatgtaacttt
caaaaatgcaagaagcacagcaaatctaataaacttaccacctccttcaataaaaagcgagaacctcct
gggagaatttaagcaccattagcagacacatcttagagc

Figure 2A SEQ ID NO: 5

MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKAMPSAGTLPWVQ
GIICNANNPCFRYPTPGEAPGVVGNFNKSIVARLFSDDARRLLYSQKDTSMKDMRKVLRRTLQQIKKSSSNL
KLQDFLVDNETFSGFLYHNLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSE
LCGLPREKLAAAERVLRSNMDILKPIRLTLNSTSPFPKELAEATKLLHSLGTLAQELFSMRSWSDMRQE
VMFLTNVNSSSSSTQIYQAVSRIVCGHPEGGLKIKSLNWEEDNNYKALFGGNGTEEDAETFYDNSTTPYC
NDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPATROVMAEVNKTQELAVFHDLEGMWEEELSPKIWTF
MENSQEMDLVRMLLDSRDNDHFWEQQLDGLDWTADQDIVAFLAKHPEDVQSSNGSVYTWREAFNETNQAIRT
ISRFMECVNLNKLEPIATEVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVKYKIRMDIDNVERTNK
IKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQMPYPCYVDDIFLRVMSRS
MPLFMTLAWIYSVAVIIKGIYVEKEARLKETMRIMGLDNSILWFSWFISSLIPLLVSAGLLVILKLGNNL
PYSDPSVVFVFLSVFAVVTILQCFLISTLFSRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIIFASL
LSPVAFGFGCEYFALFEEQIGVQWDNLFESPVEEDGFNLTTSVSMMLFDTFLYGVMTWYIEAVFPGQYGI
PRPWYFPCTKSYWFGESDEKSHPGSNQKRISICMEEETHLKLGVSIQNLVKVYRDGMKVAVDGLALNF
YEQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQNLGVCPQHNVLFDMLTVEEHI
WFIARLKLSEKHVKAEMEQLADVGLPSSKLKSKTSQLSGGMQRKLSVALAFVGGSKVILDEPTAGVDP
YSRRGIWELLLKYRQGRTIILSTHMDVGLDRIAIISHGKLCCVGSSFLKNQLGTGYLLTLVKKDVE
SSLSSCRNSSSTVSYLKKEDSVSQSSDAGLGDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHELTYV
LPYEAKEGAFVELFHEIDRLSDLGISYGISETTLEEIFLKVAEESGVDAETSDGTLPARNRRAFGDK
QSLRPFTEDDAADPNDSIDPESRETDLLSGMDGKGSYQVKGWKLQQQFVALLWKRLLIARRSRKGFFA
QIVLPAVFVCIALVFLIVPPFGKYPSELQPMWYNEQYTFVSNDAPEDTGTLELLNALTDPGFGTRCME
GNPIPDTPCQAGEEWTAPVPQTIMDLFQNGNWTMNPSPACQCSSDKIKMLPVCPPGAGGLPPPQRKQ
NTADILQDLTGRNISDYLKTYVQIIAKSLKNKIWNNEFRYGGFSLGVSNTQALPPSQEVNDAIKQMKKHL
KLAkdSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVINNAILRANLQGENPSHYGITAFN
HPLNLTKQQLSEVALMTTSVDVLSICVIFAMSFVPASFVFLIQERVSKAKHLQFISGVKPVYIYWLNSFV
WDMCNYVVPATLVIIIFICFQQKSYVSSTNLPVLALLLLLYGWSITPLMPASFVFKIPSTAYVVLTSVNL
FIGINGSVATFVLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSW
DLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRIIDGGGQNDILEIKEL
TKIYRRKRKPAVDRIICVGIPPGECFGLLVNGAGKSSTFKMLTGDTTVTRGDAFLNKNLSILSNIHEVHQNM
GYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKVGWEAIRKLGLVKYGEKYAGNYSGGNKRKLSTAMAL
IGGPPVVFLDEPTTGMDPKARRFLWNCALSVVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQH
LKNRFGDGYTIVVRIAGSNPDLKPVDQDFGLAFPGSVLKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRL
HIEDYSVSQTTLQVVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV*

Figure 2B SEQ ID NO: 6

GTCCCTGCTGTGAGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGGGTG
CTGGCTGAGGGAACATGGCTTGTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCA
CTTTCAGAAGAAGACAAACATGTCAGCTGTTACTGGAAGTGGCCTGGCCTCTATTTATCT
TCCTGATCCTGATCTCTGTTTCGGCTGAGCTACCCACCCTATGAACAACATGAATGCCATT
TTCCAAATAAAGCCATGCCCTCTGCAGGAACACTTCCTTGGGTTCAGGGGATTATCTGTA
ATGCCAACAACCCCTGTTTCCGTTACCCGACTCCTGGGGAGGCTCCCGAGTTGTTGGAA
ACTTTAACAATCCATTGTGGCTCGCCTGTTCTCAGATGCTCGGAGGCTTCTTTTATACA
GCCAGAAAGACACCAGCATGAAGGACATGCGCAAAGTTCTGAGAACATTACAGCAGATCA
AGAAATCCAGCTCAAACCTGAAGCTTCAAGATTTCTGGTGGACAATGAAACCTTCTCTG
GGTTCCTGTATCACAACTCTCTCTCCCAAAGTCTACTGTGGACAAGATGCTGAGGGCTG
ATGTCATTCTCCACAAGGTATTTTTGCAAGGCTACCAGTTACATTTGACAAGTCTGTGCA
ATGGATCAAAATCAGAAGAGATGATTCAACTTGGTGACCAAGAAGTTTCTGAGCTTTGTG
GCCTACCAAGGGAGAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCC
TGAAGCCAATCCTGAGAACACTAACTCTACATCTCCCTTCCCGAGCAAGGAGCTGGCTG
AAGCCACAAAAACATTGCTGCATAGTCTTGGGACTCTGGCCCAGGAGCTGTTTCAGCATGA
GAAGCTGGAGTGACATGCGACAGGAGGTGATGTTTCTGACCAATGTGAACAGCTCCAGCT
CCTCCACCCAAATCTACCAGGCTGTGTCTCGTATTGTCTGCGGGCATCCCGAGGGAGGGG
GGCTGAAGATCAAGTCTCTCAACTGGTATGAGGACAACAATAACAAAGCCCTCTTTGGAG
GCAATGGCACTGAGGAAGATGCTGAAACCTTCTATGACAACCTCTACAACCTCCTTACTGCA
ATGATTTGATGAAGAATTTGGAGTCTAGTCCTCTTTCCCGCATTATCTGGAAAGCTCTGA
AGCCGCTGCTCGTTGGGAAGATCCTGTATACACCTGACACTCCAGCCACAAGGCAGGTCA
TGGCTGAGGTGAACAAGACCTTCCAGGAACCTGGCTGTGTTCCATGATCTGGAAGGCATGT
GGGAGGAACTCAGCCCCAAGATCTGGACCTTCATGGAGAACAGCCAAGAAATGGACCTTG
TCCGGATGCTGTTGGACAGCAGGGACAATGACCACTTTTGGGAACAGCAGTTGGATGGCT
TAGATTGGACAGCCCAAGACATCGTGGCGTTTTTGGCCAAGCACCCAGAGGATGTCCAGT
CCAGTAATGGTTCTGTGTACACCTGGAGAGAAGCTTTCAACGAGACTAACCAGGCAATCC
GGACCATATCTCGCTTCATGGAGTGTGTCAACCTGAACAAGCTAGAACCCATAGCAACAG
AAGTCTGGCTCATCAACAAGTCCATGGAGCTGCTGGATGAGAGGAAGTTCTGGGCTGGTA
TTGTGTTCACTGGAATTACTCCAGGCAGCATTGAGCTGCCCCATCATGTCAAGTACAAGA
TCCGAATGGACATTGACAATGTGGAGAGGACAAATAAAATCAAGGATGGGTACTGGGACC
CTGGTCCTCGAGCTGACCCCTTTGAGGACATGCGGTACGTCTGGGGGGGCTTCGCCTACT
TGCAGGATGTGGTGGAGCAGGCAATCATCAGGGTCTGACGGGCACCGAGAAGAAAACCTG

GTGTCTATATGCAACAGATGCCCTATCCCTGTTACGTTGATGACATCTTTCTGCGGGTGA
TGAGCCGGTCAATGCCCCCTCTTCATGACGCTGGCCTGGATTTACTCAGTGGCTGTGATCA
TCAAGGGCATCGTGTATGAGAAGGAGGCACGGCTGAAAGAGACCATGCGGATCATGGGCC
TGGACAACAGCATCCTCTGGTTTAGCTGGTTCATTAGTAGCCTCATTCTCTTCTGTGA
GCGCTGGCCTGCTAGTGGTCATCCTGAAGTTAGGAAACCTGCTGCCCTACAGTGATCCCA
GCGTGGTGTGTTGTCTTCTGTCCGTGTTTGCTGTGGTGACAATCCTGCAGTGCTTCCTGA
TTAGCACACTCTTCTCCAGAGCCAACCTGGCAGCAGCCTGTGGGGGCATCATCTACTTCA
CGCTGTACCTGCCCTACGTCTGTGTGTGGCATGGCAGGACTACGTGGGCTTCACACTCA
AGATCTTCGCTAGCCTGCTGTCTCCTGTGGCTTTTGGGTTTGGCTGTGAGTACTTTGCCC
TTTTTGAGGAGCAGGGCATTGGAGTGCAGTGGGACAACCTGTTTGAGAGTCCTGTGGAGG
AAGATGGCTTCAATCTCACCATTCTGGTCTCCATGATGCTGTTTGACACCTTCCTCTATG
GGGTGATGACCTGGTACATTGAGGCTGTCTTTCCAGGCCAGTACGGAATTCACAGGCCCT
GGTATTTTCTTGCACCAAGTCTTACTGGTTTGGCGAGGAAAGTGATGAGAAGAGCCACC
CTGGTTCCAACAGAGAATATCAGAAATCTGCATGGAGGAGGAACCCACCCACTTGA
AGCTGGGCGTGTCCATTGAGAACCCTGGTAAAAGTCTACCGAGATGGGATGAAGGTGGCTG
TCGATGGCCTGGCAGTGAATTTTTATGAGGGCCAGATCACCTCCTTCCTGGGCCACAATG
GAGCGGGGAAGACGACCACCATGTCAATCCTGACCGGGTTGTTCCCCCGACCTCGGGCA
CCGCCTACATCCTGGGAAAAGACATTCTGCTCTGAGATGAGCACCATCCGGCAGAACCTGG
GGGTCTGTCCCCAGCATAACGTGCTGTTTGACATGCTGACTGTGGAAGAACACATCTGGT
TCTATGCCCCTTGAAAGGGCTCTCTGAGAAGCACGTGAAGGCGGAGATGGAGCAGATGG
CCCTGGATGTTGGTTTGCCATCAAGCAAGCTGAAAAGCAAAACAAGCCAGCTGTCAGGTG
GAATGCAGAGAAAGCTATCTGTGGCCTTGGCCTTTGTGCGGGGATCTAAGGTGTCTATC
TGGATGAACCCACAGCTGGTGTGGACCCTTACTCCCGCAGGGGAATATGGGAGCTGCTGC
TGAAATACCGACAAGGCCGACCATTTATCTCTCTACACACCACATGGATGAAGCGGACG
TCCTGGGGGACAGGATTGCCATCATCTCCCATGGGAAGCTGTGCTGTGTGGGCTCCTCCC
TGTTTCTGAAGAACCAGCTGGGAACAGGCTACTACCTGACCTTGGTCAAGAAAGATGTGG
AATCCTCCCTCAGTTCTGCAGAAACAGTAGTAGCACTGTGTCATACCTGAAAAAGGAGG
ACAGTGTTTCTCAGAGCAGTTCTGATGCTGGCCTGGGCAGCGACCATGAGAGTGACACGC
TGACCATCGATGTCTCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCCCCGC
TGGTGGAAAGACATAGGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGG
GAGCCTTTGTGGAACCTTTTCATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTA
GTTATGGCATCTCAGAGACGACCCTGGAAGAAATATTCCTCAAGGTGGCCGAAGAGAGTG

GGGTGGATGCTGAGACCTCAGATGGTACCTTGCCAGCAAGACGAAACAGGCGGGCCTTCG
GGGACAAGCAGAGCTGTCTTCGCCCCGTTCACTGAAGATGATGCTGCTGATCCAAATGATT
CTGACATAGACCCAGAATCCAGAGAGACAGACTTGCTCAGTGGGATGGATGGCAAAGGGT
CCTACCAGGTGAAAGGCTGGAACTTACACAGCAACAGTTTGTGGCCCTTTTGTGGAAGA
GACTGCTAATTGCCAGACGGAGTCGGAAAGGATTTTTTGTCTCAGATTGTCTTGCCAGCTG
TGTTTGTCTGCATTGCCCTTGTGTTTACAGCTGATCGTGCCACCCTTTGGCAAGTACCCCA
GCCTGGAACCTTCAGCCCTGGATGTACAACGAACAGTACACATTTGTCTCAGCAATGATGCTC
CTGAGGACACGGGAACCTTGGAACCTTAAACGCCCTCACCAAAGACCCTGGCTTCGGGA
CCCGCTGTATGGAAGGAAACCAATCCCAGACACGCCCTGCCAGGCAGGGGAGGAAGAGT
GGACCACTGCCCCAGTTCCCCAGACCATCATGGACCTCTTCCAGAATGGGAACTGGACAA
TGCAGAACCCTTCACCTGCATGCCAGTGTAGCAGCGACAAAATCAAGAAGATGCTGCCTG
TGTGTCCCCCAGGGGCAGGGGGGCTGCCTCCTCCACAAAGAAAACAAAACACTGCAGATA
TCCTTCAGGACCTGACAGGAAGAAACATTTTCGGATTATCTGGTGAAGACGTATGTGCAGA
TCATAGCCAAAAGCTTAAAGAACAAGATCTGGGTGAATGAGTTTAGGTATGGCGGCTTTT
CCCTGGGTGTCTAGTAATACTCAAGCACTTCTCCGAGTCAAGAAGTTAATGATGCCATCA
AACAAATGAAGAAACACCTAAAGCTGGCCAAGGACAGTTCTGCAGATCGATTTCTCAACA
GCTTGGGAAGATTTATGACAGGACTGGACACCAGAAATAATGTCAAGGTGTGGTTCAATA
ACAAGGGCTGGCATGCAATCAGCTCTTTCTGAATGTCATCAACAATGCCATTCTCCGGG
CCAACCTGCAAAAGGGAGAGAACCCTAGCCATTATGGAATTACTGCTTTCAATCATCCCC
TGAATCTCACCAAGCAGCAGCTCTCAGAGGTGGCTCTGATGACCACATCAGTGGATGTCC
TTGTGTCCATCTGTGTCATCTTTGCAATGTCTTCGTCCCAGCCAGCTTTGTCTGATTCC
TGATCCAGGAGCGGGTCAGCAAAAGCAAAACACCTGCAGTTTCATCAGTGGAGTGAAGCCTG
TCATCTACTGGCTCTCTAATTTTGTCTGGGATATGTGCAATTACGTTGTCCCTGCCACAC
TGGTCATTATCATCTTCATCTGCTTCCAGCAGAAGTCTATGTGTCCTCCACCAATCTGC
CTGTGCTAGCCCTTCTACTTTTGTCTGTATGGGTGGTCAATCACACCTCTCATGTACCCAG
CCTCCTTTGTGTTCAAGATCCCCAGCACAGCCTATGTGGTGTCTACCAGCGTGAACCTCT
TCATTGGCATTAAATGGCAGCGTGGCCACCTTTGTGCTGGAGCTGTTACCGACAATAAGC
TGAATAATATCAATGATATCCTGAAGTCCGTGTTCTTGATCTTCCCACATTTTGCCTGG
GACGAGGGCTCATCGACATGGTGAAAAACCAGGCAATGGCTGATGCCCTGGAAAGGTTTG
GGGAGAATCGCTTTGTGTCAACATTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCCA
TGGCCGTGGAAGGGGTGGTGTCTTCCCTCATTACTGTTCTGATCCAGTACAGATTCTTCA
TCAGGCCCAGACCTGTAAATGCAAAGCTATCTCCTCTGAATGATGAAGATGAAGATGTGA
GGCGGGAAAGACAGAGAAATTCTTGATGGTGGAGGCCAGAATGACATCTTAGAAATCAAGG
AGTTGACGAAGATATATAGAAGGAAGCGGAAGCCTGCTGTTGACAGGATTTGCGTGGGCA

TTCCTCCTGGTGAGTGCTTTGGGCTCCTGGGAGTTAATGGGGCTGGAAAATCATCAACTT
TCAAGATGTTAACAGGAGATACCACTGTTACCAGAGGAGATGCTTTCCTTAACAAAAATA
GTATCTTATCAAACATCCATGAAGTACATCAGAACATGGGCTACTGCCCTCAGTTTGATG
CCATCACAGAGCTGTTGACTGGGAGAGAACACGTGGAGTTCTTTGCCCTTTTGAGAGGAG
TCCCAGAGAAAGAAGTTGGCAAGGTTGGTGAGTGGGCGATTTCGGAACTGGGCCTCGTGA
AGTATGGAGAAAAATATGCTGGTAACTATAGTGGAGGCAACAAACGCAAGCTCTCTACAG
CCATGGCTTTTGATCGGCGGGCCTCCTGTGGTGTCTTGATGAACCCACCACAGGCATGG
ATCCCAAAGCCCGGCGGTTCTTGTGGAATTGTGCCCTAAGTGTGTCAAGGAGGGGAGAT
CAGTAGTGCTTACATCTCATAGTATGGAAGAATGTGAAGCTCTTTCAGCTAGGATGGCAA
TCATGGTCAATGGAAGGTTTCAGGTGCCTTGGCAGTGTCCAGCATCTAAAAATAGGTTTG
GAGATGGTTATACAATAGTTGTACGAATAGCAGGGTCCAACCCGGACCTGAAGCCTGTCC
AGGATTTCTTTGGACTTGCATTTCTGGAAGTGTCTAAAAGAGAAACACCGGAACATGC
TACAATACCAGCTTCCATCTTCATTATCTTCTCTGGCCAGGATATTTCAGCATCCTCTCCC
AGAGCAAAAAGCGACTCCACATAGAAGACTACTCTGTTTCTCAGACAACACTTGACCAAG
TATTTGTGAACTTTTGCCAAGGACCAAAGTGATGATGACCACTTAAAAGACCTCTCATTAC
ACAAAAACCAGACAGTAGTGGACGTTGCAGTTCTCACATCTTTTCTACAGGATGAGAAAG
TGAAAGAAAGCTATGTATGAAGAATCCTGTTTCATACGGGGTGGCTGAAAGTAAAGAGGAA
CTAGACTTTCTTTTGACCATGTGAAGTGTGTGGAGAAAAGAGCCAGAAGTTGATGTGG
GAAGAAGTAACTGGATACTGTACTGATACTATTCAATGCAATGCAATTCAATGCAATGA
AAACAAAATTCCATTACAGGGGCAGTGCCTTTGTAGCCTATGTCTTGTATGGCTCTCAAG
TGAAAGACTTGAATTTAGTTTTTTTACCTATACCTATGTGAACTCTATTATGGAACCCAA
TGGACATATGGGTTTGAACTCACACTTTTTTTTTTTTTTTTTTTGTTCTGTGTATTCTCATT
GGGGTTGCAACAATAATTCATCAAGTAATCATGGCCAGCGATTATTGATCAAAATCAAAA
GGTAATGCACATCCTCATTCACTAAGCCATGCCATGCCCAGGAGACTGGTTTTCCCGGTGA
CACATCCATTGCTGGCAATGAGTGTGCCAGAGTTATTAGTGCCAAGTTTTTCAGAAAGTT
TGAAGCACCATGGTGTGTCTATGCTCACTTTTGTGAAAGCTGCTCTGCTCAGAGTCTATCA
ACATTGAATATCAGTTGACAGAATGGTGCCATGCGTGGCTAACATCCTGCTTTGATTCCC
TCTGATAAGCTGTTCTGGTGGCAGTAACATGCAACAAAAATGTGGGTGTCTCCAGGCACG
GGAAACTTGGTTCCATTGTTATATTGTCCTATGCTTCGAGCCATGGGTCTACAGGGTCAT
CCTTATGAGACTCTTAAATATACTTAGATCCTGGTAAGAGGCAAAGAATCAACAGCCAAA
CTGCTGGGGCTGCAACTGCTGAAGCCAGGGCATGGGATTAAAGAGATTGTGCGTTCAAAC
CTAGGGAAGCCTGTGCCCATTGTCTGACTGTCTGCTAACATGGTACACTGCATCTCAA
GATGTTTATCTGACACAAGTGTATTATTCTGGCTTTTTTGAATTAATCTAGAAAATGAAA

Figure 3

Promoter, 8797 bp

Distances numbered using first base of promoter as 1

Name	Pos. of 1st base in sense strand	Hit Site	% Match	Strand	SEQ ID No.
LXRE		Target: AGGTCA (NNNN){AGGTCA			7
DR4	-7531	AGAGGCAGGTGGATCATTTGAGGTCA	88	sense	8
DR4	-5085	TTGAGCGGGTGATCACTTGAGGTCA	88	antisense	9
DR4	-4389	CAAGCGGGCAGATCACTTGAGGTTA	88	antisense	10
DR4	-1641	CAAGTGGGCAGCTCACCTCAGGTCA	94	antisense	11
DR1		None			
PPAR		Target: NNNNN(A)NN(T)TGACCT(N/NN)TGACCT			12
DR2	-7718	CTTTGA (A)GC (C) TGATCATATGACCT	88	antisense	13
DR2	-7521	AGGCTG (G) TC (T) CGAACTCCTGACCT	88	antisense	14
DR2	-5708	CTTAAT (T) GG (T) GGWGTGTTGACCT	91	antisense	15
DR2	-2894	CAGGAT (G) GC (G) TAAACTCCTGACCT	88	antisense	16
DR2	-1649	AGGTTG (G) TT (T) CGAACTCCTGACCT	88	sense	17
DR2	-1140	TCAAGG (T) AG (G) AGACCTTGTGGCCT	88	sense	18
DR1		None			

Name	Pos. of 1st base in sense strand	Hit Site	% Match	Strand
SREBP		Target: ATCACCCCAC		19
	-8523	GAGATGTGCTATGACCCCAC	90	antisense
	-3651	GTGAGCCCAGATCACACCAC	90	antisense
	-7747	TCCATCCATCCACACCCCAC	80	antisense
	-5485	CCCTTTTATTAAACAGCTCAC	80	antisense
	-5248	GTAAGCCCAAGATCATGCCCAC	80	antisense
	-5073	ACCTCAAGTGATCACCCGCC	80	sense
	-2252	GGCTCAAGCGATCCTGCCAC	80	antisense
	-2209	CCATGATTGGATCATCTGCAC	80	sense
	-1794	GTGAGTCGAGATCATGCCCAC	80	antisense
	-519	TGCTTTTGTTTTCCCCCCAC	80	antisense
	-478	CCGCCTTCCCTTCACCCCAG	80	sense
	-158	ACCCTCCACCCCACCCCCAC	80	sense
				31

ROR	Target: (w) {0,8}WRGGTCA	
32		
33	-8435 CTGGCAAGGATGGGTCA	100 sense
34	-8434 TGGCAAGGATGGGTCA	100 sense
35	-7025 AAAAAGCACCAAGGTCA	100 antisense
36	-3989 AGAAGATGCCAGGGTCA	100 sense
37	-2638 GAGGAGATGGAGGGTCA	100 sense

Exon 1, 303 bp

Distances numbered using start of Exon 1 as + 1

Name	Pos. of 1st base in +	Hit Site	% Match	Strand	Q ID NO.
LXRE		Target: AGGTCA (NNNN)AGGTCA			7
DR4	4	CCGAGCGCAGAGGTTACTATCGGTCA	92	antisense	38
DR1		None			

PPAR

DR2

DR1

SREBP

ROR

5' Intron 1, 930 bp

Positions numbered using the first position in intron 1 as + 1

Name	Pos. of 1st base in +	Hit Site	Matc	Strand	SEQ ID N
LXRE		Target: AGGTCA (NNNN)AGGTCA			7
DR4	458	GCCAATTCCCAAGTCAAGACAGACCA	88	antisense	39
DR1		None			
PPAR		Target: NNNNNN (A) NN (T) TGACCT (N/NN) TGACCT			12
DR2		None			
DR1		None			
SREBP		Target: ATCACCCAC			19
	326	GGACCTGCAGCTCTCCCCAC	80	antisense	40
ROR		Target: (W) {0,8}WRGGTCA			32
	17	AACGCCCAAGTAAGTCA	94	antisense	41
	161	GAGCTCGTACTAGGACA	94	antisense	42
	181	GCAGAGTCTGGGTCA	94	antisense	43
	181	CGAGAGTCTGGGTCA	94	antisense	44
	478	AGCCAATTCACAGGTCA	94	antisense	45
	559	ACGGACCGTTTGGGACA	94	antisense	46
	559	CACGACCGTTTGGGACA	94	antisense	47
	559	CCACGACCGTTTGGGACA	94	antisense	48
	589	ACTAGAGGCTTGGGTCT	94	sense	49
	590	CTAGAGGCTTGGGTCT	94	sense	50
	612	CCCTACCCCTCAGGTCA	94	antisense	51
	612	TCCCTACCCCTCAGGTCA	94	antisense	52
	668	GSTCTGCCGACGAGACA	94	antisense	53
	864	TTTGTAGTGAGANGTTA	94	sense	54

3' end of Intron 1, 12504 bp

Positions numbered using the first base 5 to the start of Exon 2 as -1

Name	Pos. of 1st base in +	Hit Site	Mat	Strand
LXRE		Target: AGGTCA (NNNN)AGGTCA		7
DR4	-7188	TGAGCAGTAGATCACTTGAGGTCA	93	sense
DR4	-11050	CGAGCTGGCGGATCACCCTGAGGTCA	86	sense
DR4	-7670	AAGCTTAACAGGTTACTGAAGGCCA	86	antisense
DR4	-4696	AGAGGTGGCGGATCACCCTGAGGTCA	86	antisense
DR1		None		

PPAR		Target: NNNNN (A) NN (T) TGACCT (N/NN) TGACCT		12
DR1		None		
DR2	-10281	CTCGAT (T) TC (C) TGACCTCGTGATCC	86	antisense
DR2	-5996	CAAAAC (A) TT (G) TGCCCTTTTGAAC	86	antisense
DR2	-932	GCGCTA (G) GG (T) TGTCCTCATACCT	86	sense
DR2	-597	CTCGAT (T) TC (T) TGACCTCGTGATCC	86	sense

SREBP		Target: ATCACCACCAC		19
	-7009	GTGAGCTGAGATCACACCAC	90	sense
	-11869	TTCAGGATGATCACCACAT	80	antisense
	-11616	GGCTCAAGTGATCCTCCAC	80	antisense
	-10100	GTGAGCCGAGATCGGCCAC	80	sense
	-8584	GTGAGTTATGATCATGCCAC	80	antisense
	-5591	CCACTGTTTCAACAACCCAC	80	sense
	-4684	ACCTCAGGTGATCGGCCAC	80	sense
	-4128	AAATGTGACAAATCTCCACAC	80	antisense
	-2524	AATATAGATATCACCCTCC	80	antisense
	-1577	CCTTTATCTACCACCCAC	80	antisense

[illegible]

Non-coding Polymorphisms (SNP):									
Intron	Numbering based on Pullinger et al., 2000	Numbering based on AJ012376.1	Sequence	Change in non-coding	Coding Nucl. Change	AA change	Genotypes	Genotypes	
Promoter	(-) 191	NA	CGGGGAAGGCGACGACGCG	G to C	Not applicable	Not applicable	CG CG	CG CG	
Promoter	(-) 17	NA	CGGGGAAGGCGACGACGCG	C to G	Not applicable	Not applicable	CG CG	CG CG	
Intron 0	(-) 1163	(-) 1435	AGTATCCCTGTTTTCACGAGAA	CCCT Insertion	Not applicable	Not applicable	CG CG	CG CG	
Intron 0	(-) 1095	(-) 1367	GTGACCCACCGAGTAGGG	A to G	Not applicable	Not applicable	CG CG	CG CG	
Intron 0	(-) 1027	(-) 1299	TATGTCGTGACATGGAGCTGTT	G to A	Not applicable	Not applicable	CG CG	CG CG	
Intron 0	(-) 720	(-) 992	CCTCCGCTGCCAGGTTACGAGTT	G to A	Not applicable	Not applicable	CG CG	CG CG	
Intron 0	(-) 461	(-) 733	GAAATTAAGTATGTAAGGAAG	A to C	Not applicable	Not applicable	CG CG	CG CG	
Intron 0	(-) 362	(-) 634	CATTTCTTAGAAGAGAGGT	A to G	Not applicable	Not applicable	CG CG	CG CG	
Intron 7	(+) 2383	(+) 2383	TTTAAAGGGGKTGATTAGGA	Frequencies unknown: G and T	Not applicable	Not applicable	CG CG	CG CG	
Intron 7	(+) 3035	(+) 3035	GAAGAAATTTKTTTTTGATT	Frequencies unknown: G and T	Not applicable	Not applicable	CG CG	CG CG	
Intron 7	(-) 15	(-) 15	TCTGTCCCAATCCCTGACG	A insertion	Not applicable	Not applicable	CG CG	CG CG	
Intron 9	(-) 42	(-) 42	AGGAGCCAAACGCTCATTTG	G insertion	Not applicable	Not applicable	CG CG	CG CG	
Intron 13	(+) 24	(+) 24	AAGCCACTGTTTAAACCACT	T to A	Not applicable	Not applicable	CG CG	CG CG	
Intron 13	(-) 83	(-) 83	GCTCCCTAGCATGAGGCTC	C to T	Not applicable	Not applicable	CG CG	CG CG	
Intron 15	(-) 4-5	(-) 4-5	TTGCCCTGTTTTCACAGAGCC	CA deletion	Not applicable	Not applicable	CG CG	CG CG	
Intron 17	(+) 2000	(+) 2000	GCGCAGTGCCTGTGTCCTTA	Frequencies unknown: C and G	Not applicable	Not applicable	CG CG	CG CG	
Intron 21	(+) 118	(+) 118	CTCTCTGTTAKCACAAGAGA	Frequencies unknown: G and T	Not applicable	Not applicable	CG CG	CG CG	

Figure 5A

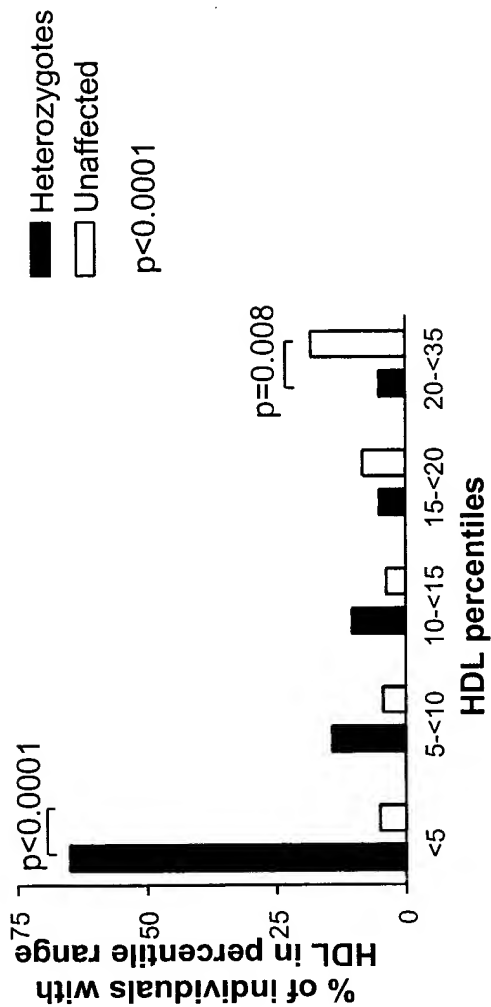


Figure 5B

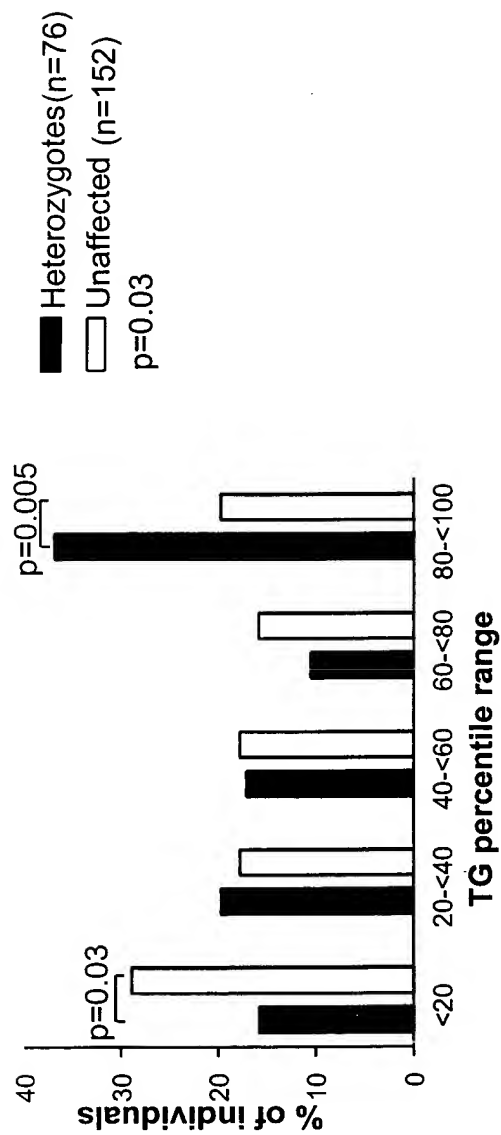


Figure 6

	TD Patients	Heterozygotes	Unaffected family members	P-value heterozygotes vs. unaffected	P-value TD patients vs. unaffected
number	5	77 ^A	156 ^A		
age (yrs) range	43.4±9.0 31-56	42.5±19.6 5-81	39.9±21.0 4-86	0.35	0.71
m/f	3/2	33/44	82/74	0.16	0.74
TC (mmol/L)	2.34±1.03	4.52±1.12	4.71±1.07	0.23	<0.0001
TG (mmol/L)	1.95±0.97	1.66±1.59	1.20±1.03	0.03	0.11
HDL (mmol/L)	0.08±0.05	0.74±0.24	1.31±0.35	<0.0001	<0.0001
LDL (mmol/L)	1.37±1.02	3.03±0.99	2.84±0.87	0.171	0.0003
ApoA-I (g/L)	0.03±0.04 (3)	0.92±0.32 (61)	1.43±0.26 (55)	<0.0001	<0.0001
ApoA-II (g/L)	0.10±0.08 (2)	0.35±0.08 (46)	0.39±0.08 (43)	0.01	<0.0001
ApoB (g/L)	0.89±0.53 (2)	0.93±0.25 (52)	0.94±0.33 (42)	0.88	0.84
CHD ≥ 20 yrs	20% (1/5)	12.9% (8/62)	4.1% (5/122)	0.03	0.10
Odds Ratio (95% CI)				3.47 (1.08-11.09)	5.85 (0.55-62.4)
Age of onset	38	48.9±8.6	60.4±12.8	0.08	

^A For TC, TG, LDL n=76 for heterozygotes, 153 for unaffected family members

Figure 7

Individual	Mutation	exon	disease (age of onset)
TD proband			
TD1	C1477R, ivs24+1G-->C 30, intron 24		CHD (38)
ABC1 heterozygotes			
TD4-201	unidentified	-	MI (<58)
FHA5-215	M1091T	22	MI (61)
FHA5-303	M1091T	22	CHD (<45)
TD1-363	C1477R	30	MI (51)
FHA3-301	Del(E,D) 1893,94	41	PVD (<54)
FHA3-305	Del(E,D) 1893,94	41	CHD (44)
FHA6-201	P2150L	48	CVA (36), fatal MI (58)
FHA2-301	R2144X	48	CAD (42), PTCA (47), femoral angioplasty (48), CABG (<50)
Unaffected family members			
FHA5-212	none	-	AP (62)
TD3-109	none	-	TIA (80)
FHA2-315	none	-	MI (51)
TD1-205	none	-	MI (62)
TD1-216	none	-	AP (47)

Figure 8

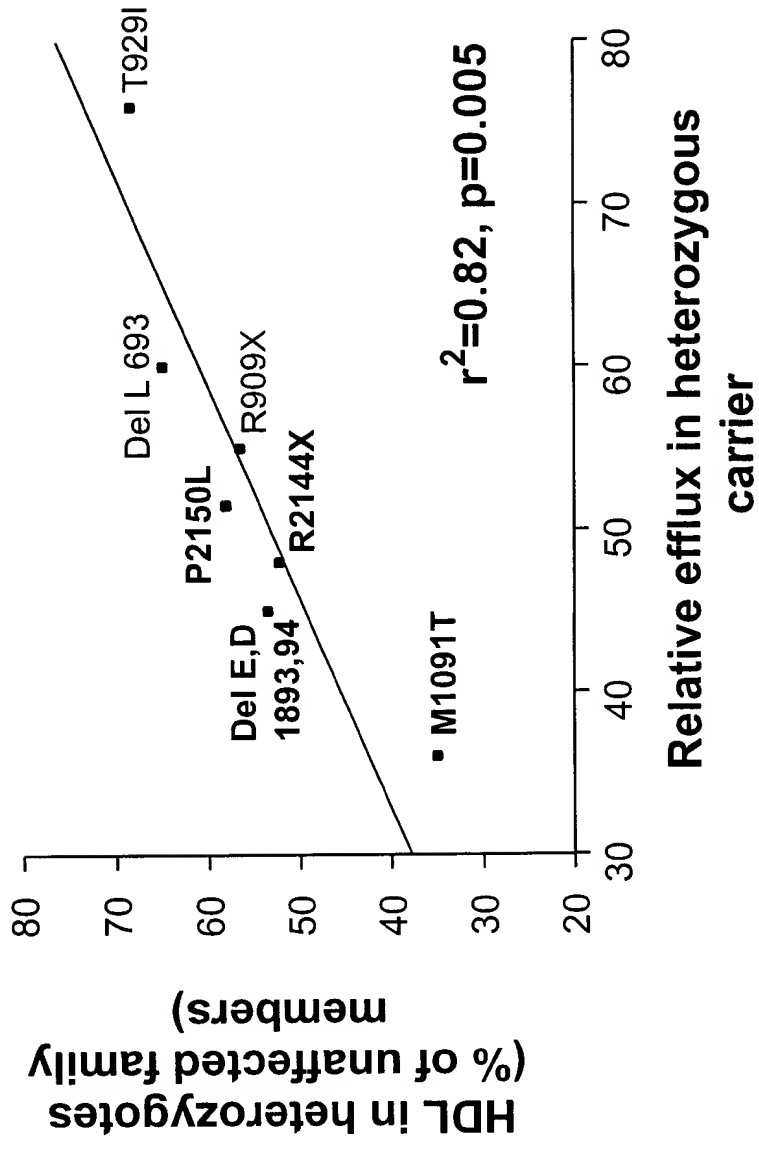


Figure 9

Figure 9

family	mutation	HDL in		HDL in unaffected		HDL in		Age and sex matched		CAD in
		heterozygotes		family members		heterozygotes		population median ^A		
		mean + SD (n)		mean + SD (n)		% of unaffected		mean + SD		
FHA1	Del L 693	0.79±0.20 (8)		1.22±0.35 (11)		64.8		1.39±0.08		-
FHA2	R2144X	0.56±0.23 (12)		1.07±0.22 (20)		52.3		1.34±0.19		+
FHA3	Del E,D 1893,94	0.77±0.24 (8)		1.44±0.38 (9)		53.5		1.30±0.17		+
FHA4	R909X	0.59±0.26 (5)		1.04±0.37 (9)		56.5		1.39±0.24		-
FHA5	M1091T	0.48±0.48 (4)		1.37±0.43 (6)		35.0		1.56±0.05		+
FHA6	P2150L	0.61±0.07 (7)		1.05 (1)		58.1		1.30±0.22		+
TD1	ivs25+1G-->C	0.78±0.06 (4)		1.35±0.29 (70)		57.8		1.22±0.22		-
TD4	del C 6825-->2145X	0.91±0.10 (2)		1.00±0.05 (3)		91.0		1.31±0.16		-
TD5	CTC6952-4TT-->2203X	0.80±0.20 (3)		1.65 (1)		48.5		1.39±0.19		-
TD1	C1477R	0.82±0.18 (9)		1.35±0.29 (70)		60.7		1.37±0.14		+
TD2	Q597R	0.82±0.07 (5)		none available		-		1.39±0.17		-
TD3	T929I	1.01±0.18 (8)		1.48±0.42 (26)		68.2		1.33±0.19		-
TD4	unidentified	0.74±0.05 (2)		1.00±0.05 (3)		73.5		1.49±0.09		+

^A Calculated based on mean the age and sex specific 50th percentile levels in the LRC population

Figure 10

	Missense	Severe	P-value	Unaffected	P-Value	P-value
	Mutations	Mutations	Missense vs.	Controls	Missense vs.	Severe vs.
	(n=33)	(n=42) ^A	Severe	(n= 156)	unaffected	unaffected
TC (mmol/L)	4.78±1.30	4.30±0.95	0.08	4.71±1.07	0.76	0.02
TG (mmol/L)	1.77±2.15	1.55±1.01	0.58	1.20±1.03	0.14	0.06
HDL (mmol/L)	0.78±0.26	0.70±0.23	0.18	1.31±0.35	<0.0001	<0.0001
LDL (mmol/L)	3.19±1.10	2.90±0.91	0.23	2.84±0.87	0.10	0.73

^A for TC, TG, LDL measurements, n=41 for severe mutations, 153 for unaffected

Figure 11

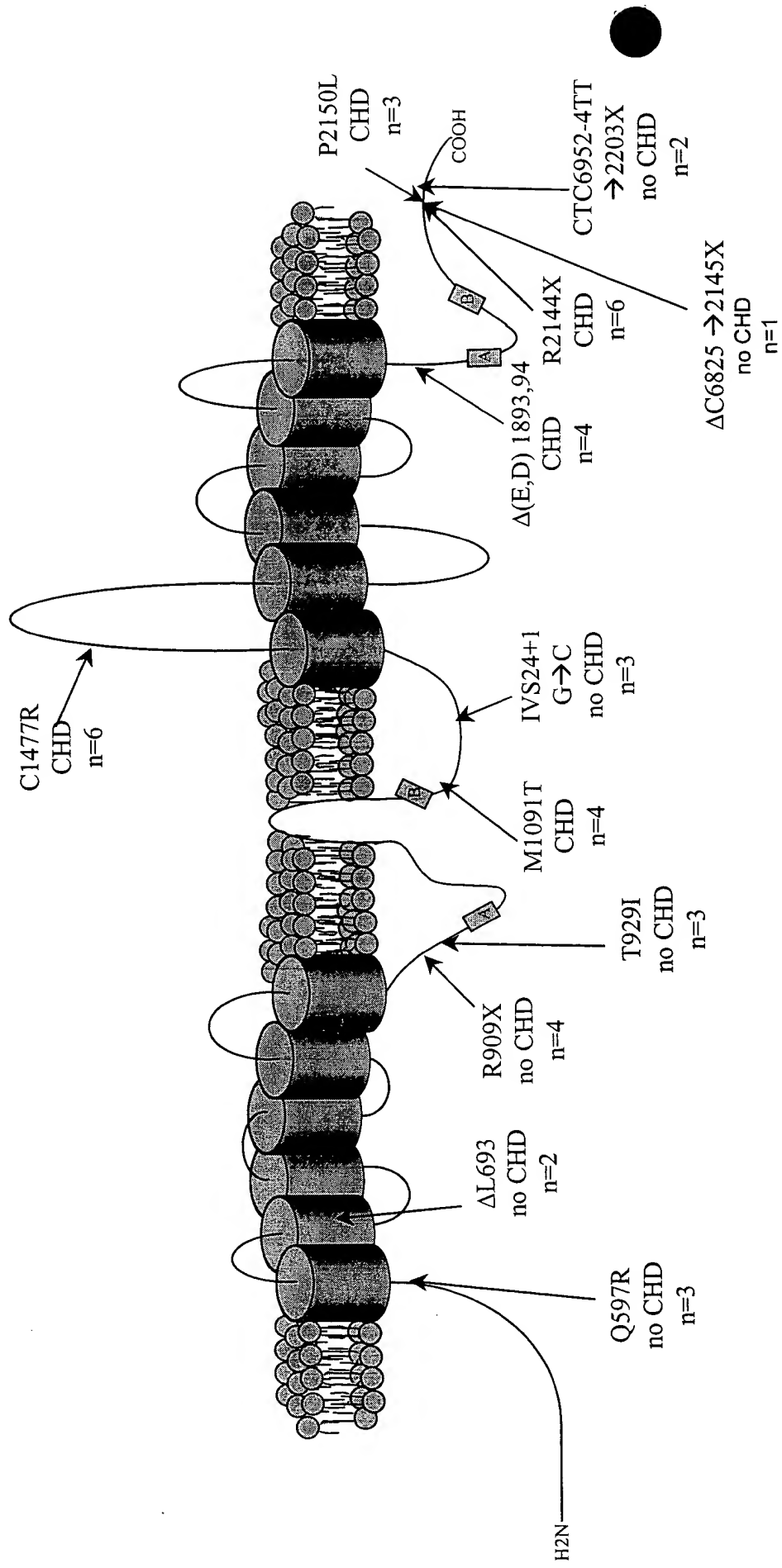


Figure 12A.

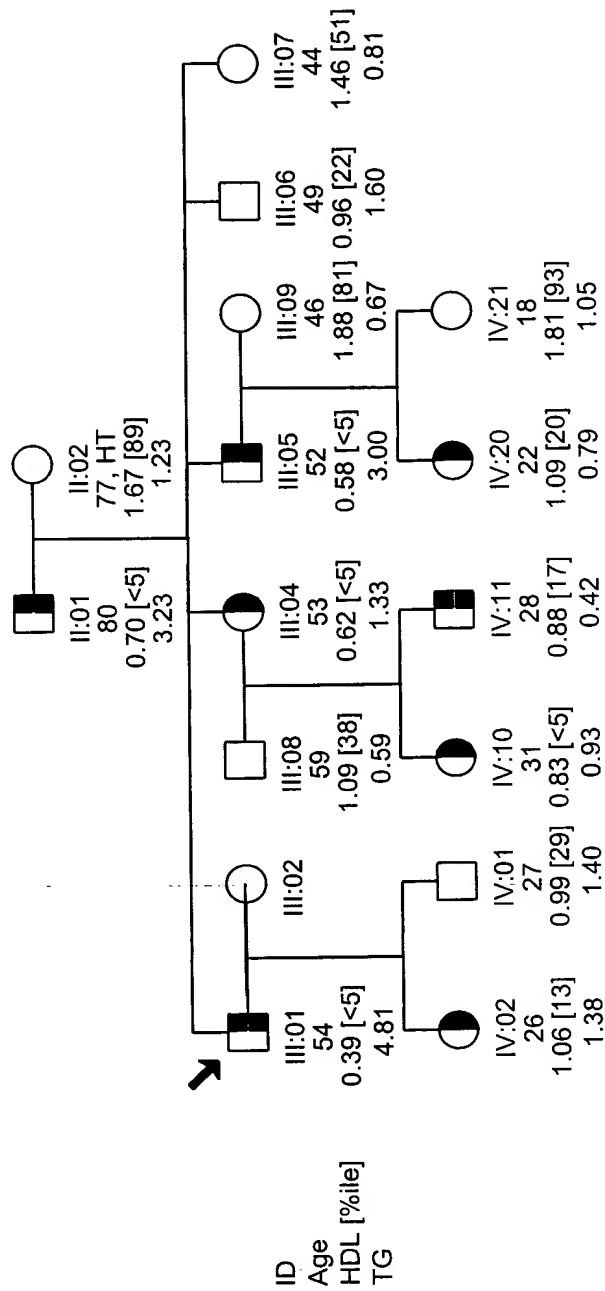


Figure 12B.

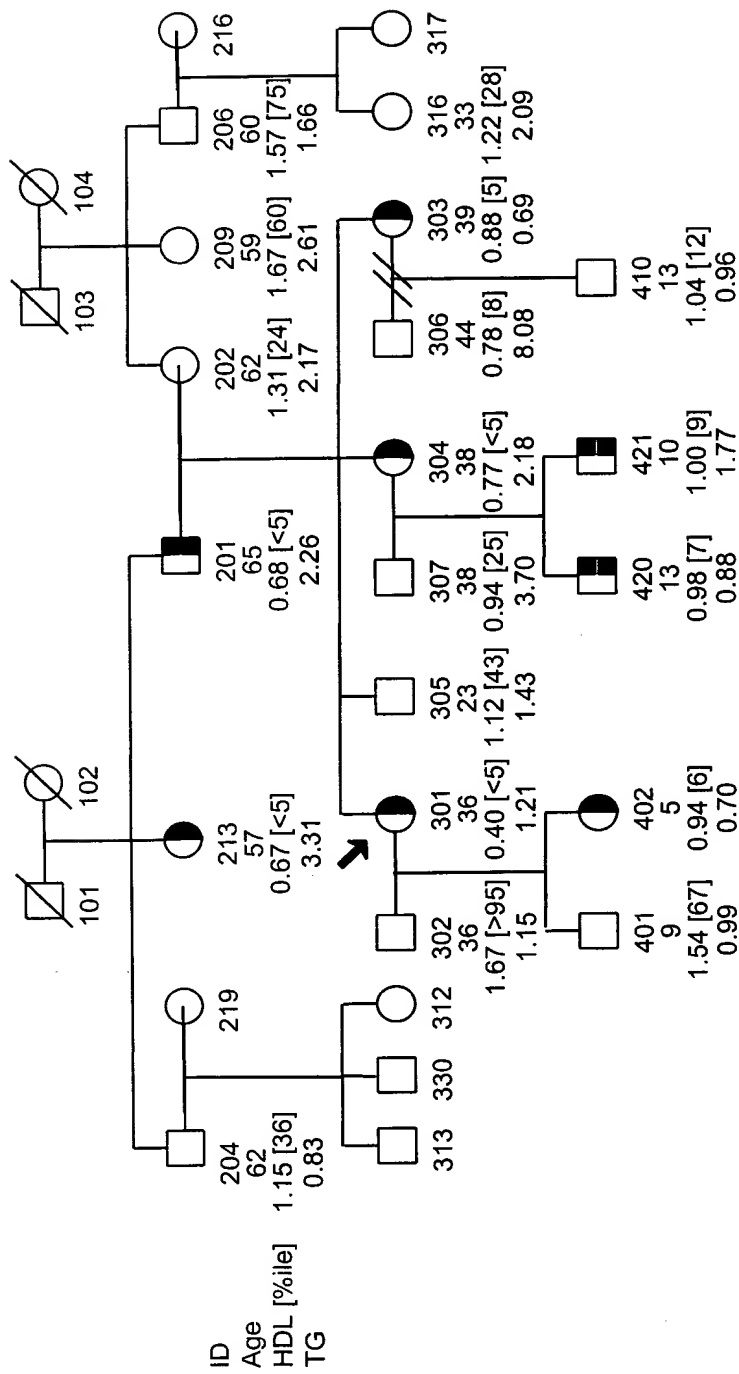


Figure 13

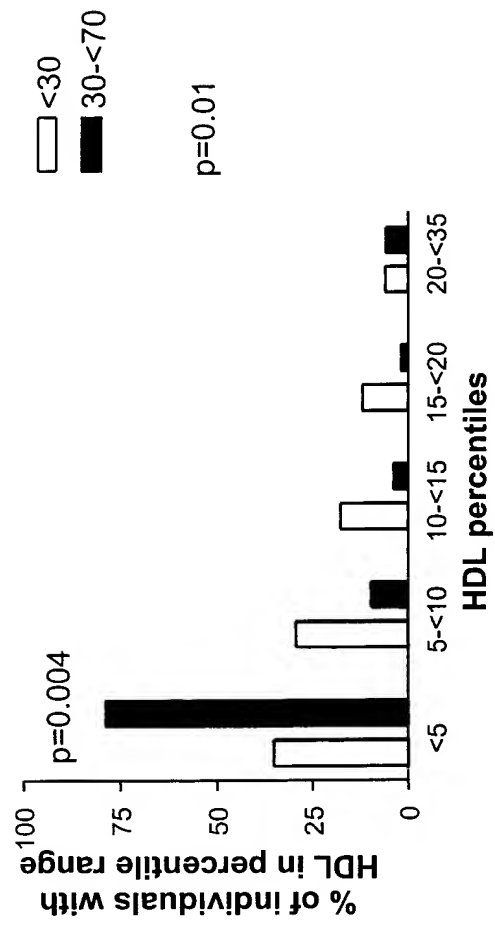


Figure 14

	Heterozygotes	Unaffected	P-value Heterozygotes
	mean±SD (n)	mean±SD (n)	vs. Unaffected
HDL (mmol/L)			
<30	0.91±0.16 (17)	1.26±0.29 (51)	<0.0001
≥30	0.66±0.24 (52)	1.32±0.36 (90)	<0.0001
Change	-0.25	+0.06	0.21
p-value <30 vs. ≥30	0.0002	0.23	
TG (mmol/L)			
<30	1.07±0.96 (16)	0.88±0.45 (51)	0.26
≥30	1.84±1.79 (52)	1.36±1.24 (87)	0.07
Change	+0.77	+0.48	0.97
p-value <30 vs. ≥30	0.03	0.001	

Figure 15A

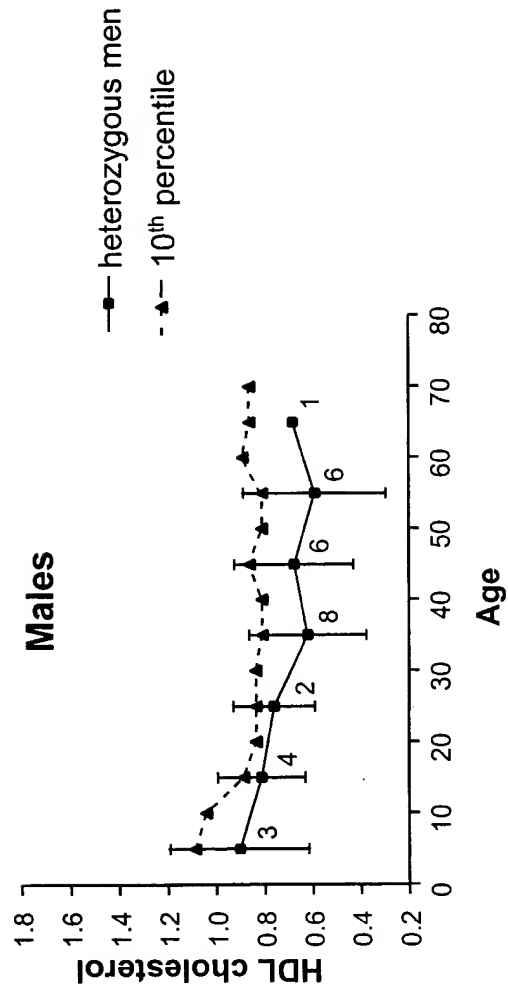


Figure 15B

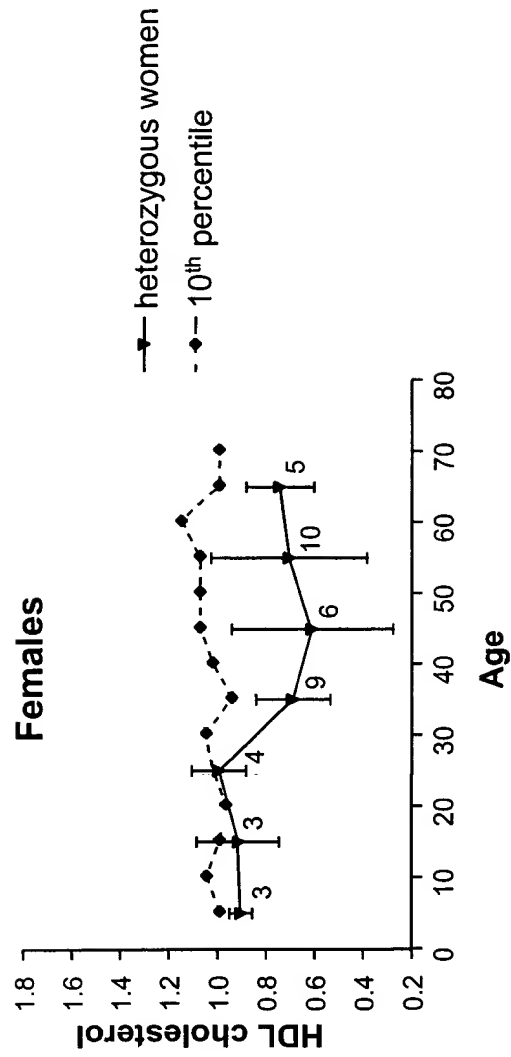


Figure 16A

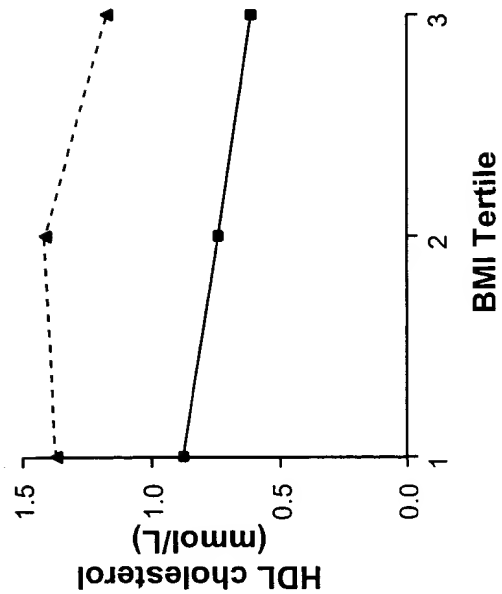


Figure 16B

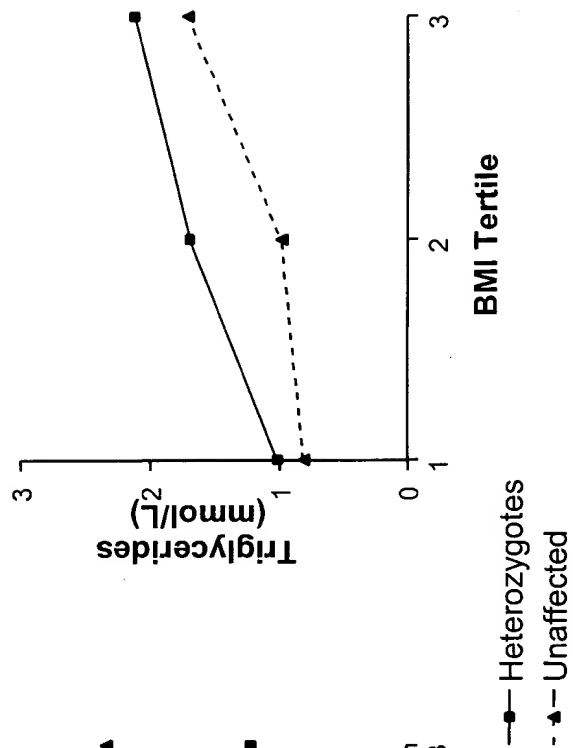


Figure 17

Variant	Amount of each dNTP (uM)	Pmol each oligo	Forward oligo (5'-->3') ^a		Annealing temp. (C)	Enzyme	Products (bp):		SEQ ID
			Reverse oligo (5'-->3') ^a				wt "A" allele	variant "B" allele	
G1051A (R219K)	187.5	20	GTATTTTTCGAAGGCTACCAAGTTACATTGACAA GATTGGCTTCAGGATGCCATGTTGGAA	60	EcoN I	177		211	
						107,70		212	
T1591C (V399A)	200	27.5	GCTGCTGTGATGGGTATCT ACCTCACTCACACCTGGGAA	57	Hph I	117,103,48,33		213	
						220,48,33		214	
G2706A (V771M)	200	27.5	CAAGTGAGTGCTTGGGATTG TGCTTTTATTACAGGGACTCCA	57	BsaA I	98, 252		215	
						350		216	
A2715C (T774P)	200	27.5	GTGATCCCAGCGTGGTGTGTCCTT GAAAGGCCAGAGGTACTCACAGCGAAGATCTTGAGGGG	55	Hph I	56,69,95		217	
						56,161		218	
G2723C (K776N)	187.5	12	TCGTTTTTATTACAGGGACTCCA CAAGTGAGTGCTTGGGATTG	55	Bgl II	269,80		219	
						349		220	
G2868A (V825I)	200	27.5	CCCATGCACTGCAGAGATTG GCAAAATTCAAATTTCTCCAGG	57	Bsa I	149, 237		221	
						386		222	
A3044G (I883M)	200	27.5	GAGAAGAGCCACCCCTGGTCCAAACCAGAAGAGGAT AAGGCAGGAGACATCGCTT	55	EcoR V	94,35		223	
						129		224	
G3911C (E1172C)	200	27.5	GAGCAGTTCTGATGCTGGCCTGGGCAGCGACCAAGCA TCTGCACCTCTCCTCCTCTG	55	BssS I	104, 37		225	
						141		226	
G5155A (R1587K)	200	27.5	CAGCTTGGGAAGATTTATGACAGGACTGGACACGA ATGCCCTGCCAACTTAC	55	BssS I	114, 31		227	
						145		228	
C5587G (S1731C)	187.5	20	GTGCAATTACGTTGTCCTGCCACACT CCATACAGCAAAAGTAGAAGGGCTAGCACACA	60	Mnl I	82,35		229	
						117		230	
G(-191)C	187.5	24	CAGCGCTTCCCGCGCGTCTTAG CCACTCACTCTCGTCCGCAATTAC	60	HgaI	287, 55, 3		233	
						342, 3		234	

C(-17)G	187.5	18	CTGCTGAGTGACTGAACACTACATAAACAGAGGCCGGG I A CCACTCACTCTCGTCCGCAATTAC	60	Rsa I	161 124, 37	235 236
C69T	187.5	24	CAGCGCTTCCGCGCGTCTTAG CCACTCACTCTCGTCCGCAATTAC	60	BsmAI	345 310,35	237 238
C127G	187.5	24	CTGGCTTTCTGCTGAGTGAC GATCAAAAGTCCCGAAACC	60	co 0109	284, 175 459	239 240
A(-362)G	187.5	24	ACTCAGTTGTATAACCCACTGAAAAAT G AGT TTCTATAGATGTTATCATCTGGG	55	Mbo II	224, 26 134, 90, 26	241 242
A(-461)C	187.5	20	ACTCAGTTGTATAACCCACTGAAAAAT G AGT TTCTATAGATGTTATCATCTGGG	55	Hinf I	150, 100 123, 100, 27	243 244
G(-720)A	187.5	20	TCATCTAAGGCACGTTGTGG CCTCAAGCCTGGAGTGACTT	60	Hpa II	450 306, 144	245 246
G(-1027)A	187.5	20	ATGGCAAAACAGTCCTCCAAG ACCCTAGCGCTGTGTCTCTG	60	Nco I	170, 41 105, 65, 41	247 248
A(-1095)G	187.5	20	ATGGCAAAACAGTCCTCCAAG ACCCTAGCGCTGTGTCTCTG	60	MspA1 I	211 172, 39	249 250
insCCCT(-1163)	187.5	20	TGTGTGTCCTCCCTTCCATT CTTGGAGGACTGTTTGCCAT	60	Mnl I	144, 28, 11, 4 100, 48, 28, 11, 4	251 252
insG319	200	27.5	CCCCCTCGCTTTATCTTTTCAGTTAATGACCAGCCCGG ATCCCCAACTCAAAAACACACA	55	Sma I	246 210, 37	253 254
G378C	200	27.5	GCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCT G CG ATCCCCAACTCAAAAACACACA	55	Acl I	108, 41, 33, 5 141, 41, 5	255 256

^a Bold indicates mismatch in oligo to create restriction site

Figure 18

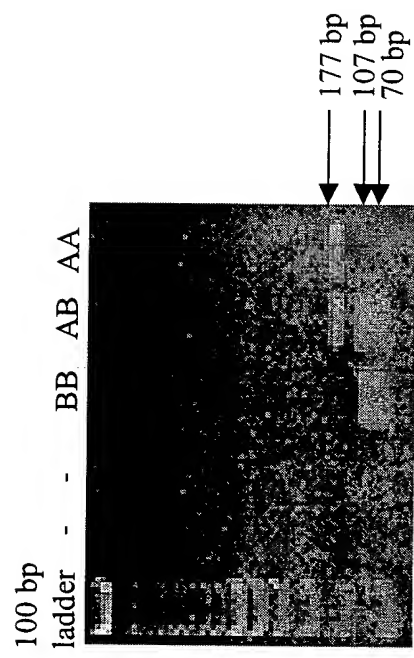


Figure 19

Nucleotide Change	Amino Acid Change	Exon	"B" allele	Frequency carrier allele (%)	N ^a	Frequency carrier allele (%)	N	Frequency carrier allele (%)	N	Frequency carrier allele (%)	N	Frequency carrier allele (%)	N	Frequency carrier allele (%)	N
Synonymous															
From Sequencing															
G869A	none	6	A	62.5	0.344	32									
C1331T	none	9	T	26.7	0.133	30									
G1343A	none	9	A	26.7	0.133	30									
T3554G	none	22	G	11.8	0.059	34									
C6842T	none	49	T	6.7	0.033	30									
Non-synonymous															
REGRESS															
G1051A	R219K	7	A	46.3	0.254	1588									
T1591C	V399A	11	C	1.6	0.008	1098									
G2706A	V771M	16	A	5.8	0.029	1270									
A2715C	T774P	16	C	0.6	0.003	1250									
G2723C	K776N	16	C	0.5	0.003	1106									
G2868A	V825I	17	A	15.7	0.081	1364									
A3044G	I883M	18	G	23.8	0.136	840									
G3911C	E1172D	24	C	5.3	0.026	1288									
G5155A	R1587K	35	A	44.3	0.259	1566									
C5587G	S1731C	38	G	0	0	558									
Dutch low HDL															
				39.9	0.227	546									
				1.2	0.006	164									
				6.3	0.031	318									
				1.3	0.006	154									
				0	0	178									
				-	-	-									
				19.7	0.102	264									
				1.3	0.007	150									
				46.1	0.271	542									
				0	0	182									
Dutch premature CAD															
				44.1	0.247	826									
				1.3	0.007	150									
				6.3	0.031	796									
				1.3	0.006	156									
				0.6	0.003	360									
				-	-	-									
				45.9	0.262	122									
				6.6	0.033	822									
				50.0	0.302	808									
				0	0	746									
Dutch control															
				47.5	0.249	686									
				1.0	0.005	416									
				3.9	0.020	406									
				0.6	0.003	326									
				1.9	0.009	318									
				-	-	-									
				26.8	0.153	314									
				8.2	0.041	440									
				51.1	0.285	446									
				0	0	158									
French Canadian															
				48.9	0.287	94									
				0	0	88									
				12.8	0.064	94									
				14.6	0.073	82									
				0	0	88									
				-	-	-									
				25.0	0.125	88									
				10.3	0.051	78									
				46.7	0.278	90									
				2.2	0.011	92									

^a N refers to the number of alleles screened.

Figure 20

R219K	AA	AB	BB	P-value		P-value	P-value
				AA vs. AB	AA vs. BB	AA vs. AB+BB	AA vs. AB+BB
n	424	330	36				
MSD	2.70±0.37	2.77±0.37	2.78±0.40	0.01	0.22	0.005	0.005
MOD	1.73±0.35	1.81±0.35	1.85±0.35	0.002	0.05	0.001	0.001
MI before trial %(n)	48.3 (205)	47.1 (155)	33.3 (12)	0.71	0.12	0.48	0.48
events during trial %(n)	17 (71)	13 (41)	11 (4)	0.10	0.49	0.09	0.09
total events ^a (%)	65.1 (276)	59.4 (196)	44.4 (16)	0.11	0.01 ^b	0.04 ^c	0.04 ^c

^a Total events is calculated as the number of events/total number of individuals. Thus, the maximum value for this variable would be 200%, as individuals may have had events both before and during the trial.

^b Odds ratio for BB vs. AA=0.43, 95% confidence interval 0.22-0.85

^c Odds ratio for AB+BB vs. AA=0.74, 95% confidence interval 0.55-0.98

Figure 21

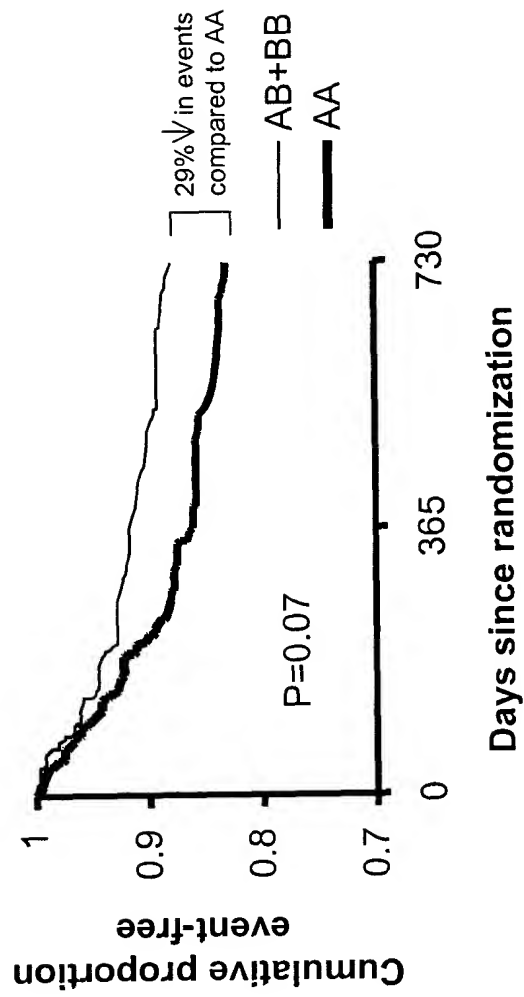


Figure 22

R219K	AA	AB	BB	P-value		
				AA vs. AB	AA vs. BB	AA vs. AB+BB
n	424	330	36			
Age	57±8	55±8	57±7	0.0007	1	0.03
BMI	25.8±2.6	26.3±2.7	25.5±2.3	0.01	0.50	0.09
Total Cholesterol	6.02±0.86	6.07±0.89	5.89±0.85	0.44	0.38	0.60
HDL Cholesterol	0.92±0.22	0.93±0.23	0.92±0.20	0.54	1	0.81
LDL Cholesterol	4.27±0.75	4.35±0.83	4.33±0.82	0.17	0.65	0.19
Triglycerides	1.84±0.77	1.78±0.78	1.42±0.49	0.29	0.001	0.08

Figure 23

		< median		> median		P-value		P-value		P-value	
		n	mean±SD	n	mean±SD	< vs. >median		AB+BB vs. AA < median		AB+BB vs. AA > median	
AB+BB											
Total Cholesterol	193	6.22±0.91	172	5.87±0.82	0.0001			0.22	0.43		
HDL cholesterol	192	0.91±0.22	171	0.94±0.23	0.21			0.12	0.37		
LDL Cholesterol	192	4.49±0.84	171	4.19±0.78	0.0005			0.03	0.57		
Triglycerides	193	1.82±0.79	172	1.65±0.72	0.03			0.02	0.85		
MSD	193	2.79±0.37	171	2.75±0.37	0.30			0.18	0.01		
MOD	193	1.83±0.36	171	1.78±0.34	0.18			0.09	0.006		
AA											
Total Cholesterol	207	6.11±0.86	217	5.94±0.84	0.04						
HDL cholesterol	206	0.88±0.20	214	0.96±0.24	0.0002						
LDL Cholesterol	205	4.32±0.77	214	4.23±0.72	0.22						
Triglycerides	206	2.02±0.82	217	1.67±0.67	<0.0001						
MSD	205	2.75±0.36	217	2.65±0.38	0.006						
MOD	205	1.77±0.34	217	1.69±0.35	0.04						

Figure 24

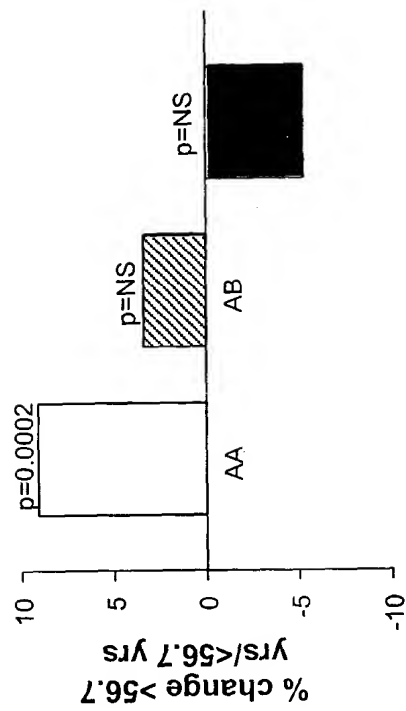


Figure 25A

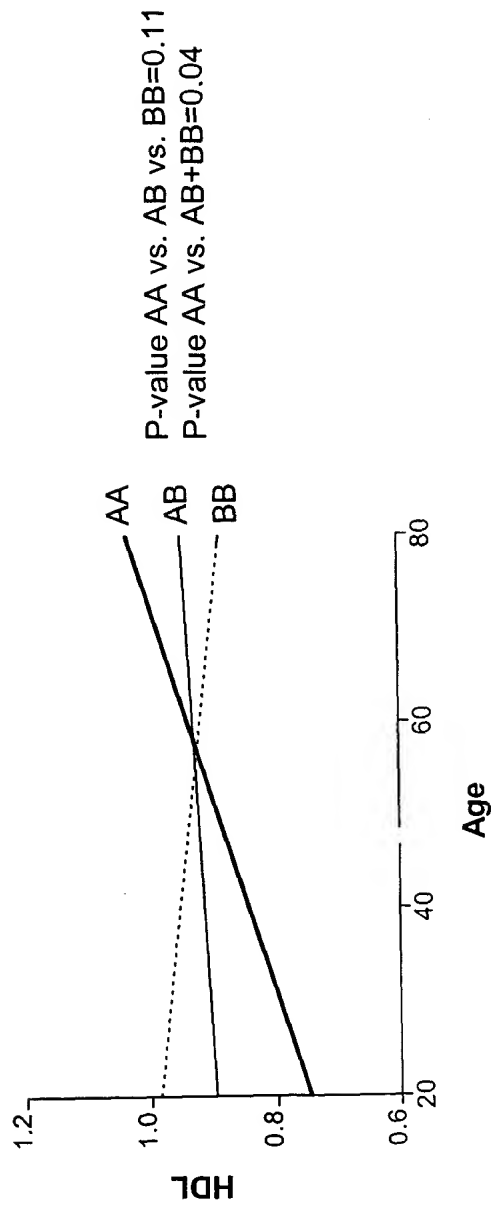


Figure 25B

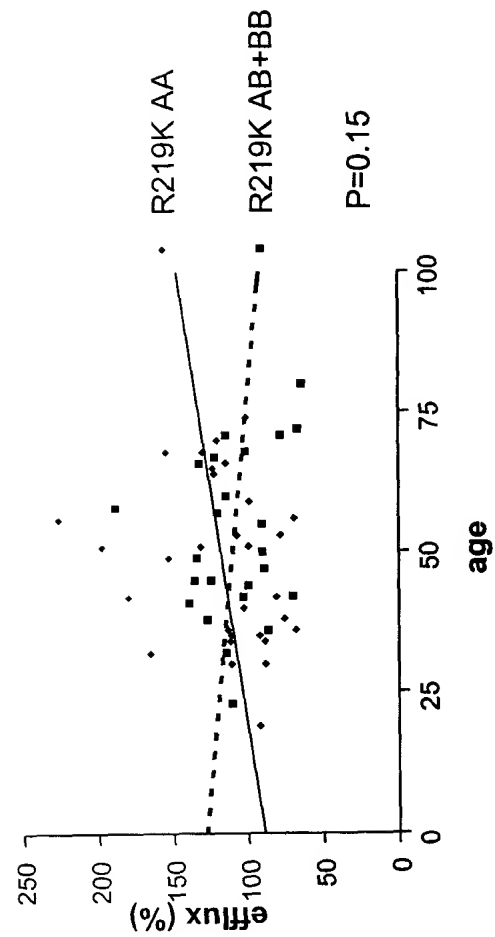


Figure 26A

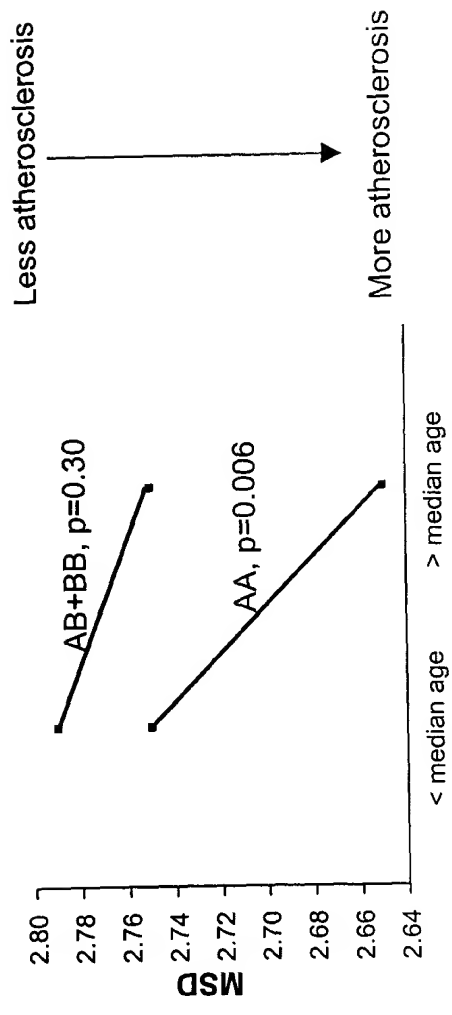


Figure 26A

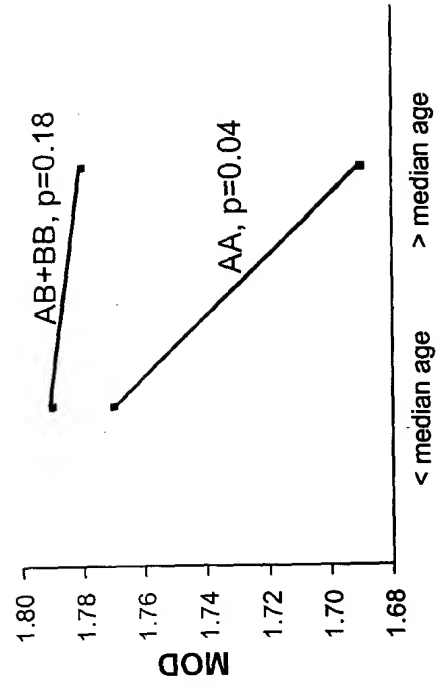


Figure 27

	South African Black ^a	Cantonese	Dutch ^b	P-value South African vs. Dutch	P-value Cantonese vs. Dutch
AA % (n)	1.3 (1)	32.7 (33)	52.5 (180)		
AB % (n)	50.7 (38)	55.4 (56)	45.2 (155)	<0.0001	<0.0001
BB % (n)	48.0 (36)	11.9 (12)	2.3 (8)		
n	75	101	343		
carrier freq.	98.67	67.33	47.52	<0.0001	0.0005
allele freq.	0.733	0.396	0.249	<0.0001	<0.0001

^a Not consistent with Hardy Weinberg equilibrium ($p=0.01$)

^b Not consistent with Hardy-Weinberg equilibrium ($p<0.001$)